ENGINEERING

~~78-0495~~

AD-A134 943

USCIPI REPORT

1

UNIVERSITY OF SOUTHERN CALIFORN

SEMIANNUAL TECHNICAL REPORT

Harry C. Andrews
Project Director

Covering Research Activity During the Period
1 April 1976 through 30 September 1976

30 September 1976

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

Sponsored by
Advanced Research Projects Agency
Contract No. F-33615-76-C-1203
ARPA Order No. 3119

IPI

IMAGE PROCESSING INSTITUTE

The views and conclusions in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U. S. Government.

USCIPI REPORT 720

SEMIANNUAL TECHNICAL REPORT

Covering Research Activity During the Period
1 April 1976 through 30 September 1976

Harry C. Andrews
Project Director
(213) 746-5514

Image Processing Institute
University of Southern California
University Park
Los Angeles, California 90007

30 September 1976

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Wright Patterson Air Force Base under Contract F-33615-76-C-1203, ARPA Order No. 3119

UNCLASSIFIED
Security Classification

AD-A134 943

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Image Processing Institute
University of Southern California, University Park
Los Angeles, California 90007

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

IMAGE PROCESSING RESEARCH

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Semiannual Technical, 1 April 1976 through 30 September 1976

5. AUTHOR(S) *(First name, middle initial, last name)*

Harry C. Andrews (Project Director)

6. REPORT DATE
30 September 1976

7a. TOTAL NO. OF PAGES
215

7b. NO. OF REFS
71

8a. CONTRACT OR GRANT NO.
F-33615-76-C-1203

b. PROJECT NO.
ARPA Order No. 3119

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

USCIPI Report 720

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Approved for release; distribution unlimited

11. SUPPLEMENTARY NOTES
----

12. SPONSORING MILITARY ACTIVITY
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

13. ABSTRACT

This technical report summarizes the image understanding and image processing research activities performed by the Image Processing Institute at the University of Southern California, during the period of 1 April 1976 to September 30, 1976 under contract number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The research program has as its primary purpose the development of techniques and systems for processing, transmitting, and analyzing images and two dimension data arrays. Four tasks are reported: Image Understanding Projects, Image Processing Projects, Smart Sensor Projects; Recent Ph. D. Dissertations. The Image Understanding project has concentrated on edge detectors, image segmentors, clustering techniques and higher level symbolic descriptors from segmented image The Image Processing projects have concentrated on a posteriori restoration, degr of freedom analyses of imaging systems, phase coding for optical image processin and artificial stereo algorithms. The Smart Sensor projects include the development of adaptive CCD image processing chips and circuitry. The Ph. D. dissertatio completed in this time period are abstracted in the next task.

****

14. Key Words: Digital Image Processing, Image Restoration, Degrees of Freedo Scene Analysis, Image Understanding, Edge Detection, Image Segmentation, CCD Arrays, CCD Processors.

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT

Security Classification

ABSTRACT

This technical report summarizes the image understanding and image processing research activities performed by the Image Processing Institute at the University of Southern California during the period of 1 April 1976 to September 30, 1976 under contract number F-33615-76-C-1203 with the Advanced Research Projects Agency Information Processing Techniques Office.

The research program has as its primary purpose the development of techniques and systems for processing, transmitting, and analyzing images and two dimensional data arrays. Four tasks are reported: Image Understanding Projects; Image Processing Projects; Smart Sensor Projects; Recent Ph.D. Dissertations. The Image Understanding project has concentrated on edge detectors, image segmentors, clustering techniques and higher level symbolic descriptors from segmented imagery. The Image Processing projects have concentrated on a posteriori restoration, degree of freedom analyses of imaging systems, phase coding for optical image processing and artificial stereo algorithms. The Smart Sensor projects include the development of adaptive CCD image processing chips and circuitry. The Ph.D. dissertations completed in this time period are abstracted in the next task.

PROJECT PARTICIPANTS

Project Director	Affiliation
Harry C. Andrews	Computer Science & Electrical Engineering
Research Staff	
Werner Frei	Electrical Engineering
Nasser E. Nahi	Electrical Engineering
Ramakant Nevatia	Computer Science
William K. Pratt	Electrical Engineering
Erica Rounds	Computer Science
Alexander A. Sawchuk	Electrical Engineering
Timothy C. Strand	Image Processing Institute

Support Staff

Marilyn Chan
Gregg Finn
Scott Johnston
Eileen Jurak
Toyone Mayeda
David Nagai
Clay Olmstead
Michael Patton
James Pepin
Ray Schmidt
Joyce Seguy
Dennis Smith
Jin Soh
Tai-ling Sung

Students

Ikram Abdou
Behnam Ashjari
Peter Chuan
Guy Coleman
Charles Hall
Chung-Kai Hsueh
Simon Lopez-Mora
Lee Martin
Dennis McCaughey
John Morton
Firouz Naderi
Javad Peyrovian

Table of Contents

1. RESEARCH PROJECT OVERVIEW

This report describes the progress and results of the University of Southern California image understanding and image processing research study for the period 1 April 1976 to 30 September 1976. The research study has been subdivided into five projects:

Image Understanding Projects

Image Processing Projects

Smart Sensors Projects

Recent Ph.D. Dissertations

Publications

The image understanding projects involve research directed toward the goal of developing generalized processing systems capable of analyzing images and extracting salient information. Specific studies include feature extraction, symbolic description, interpretation, clustering techniques, edge detectors, segmentors, and systems analysis. The image processing projects include research on novel image coding techniques based upon results of the image understanding study, advanced image restoration methods, including a posteriori phase computation techniques, degree of freedom, modelling in support of image understanding, and studies of phase coding for optical filtering techniques and artificial stereo algorithms. The smart sensors projects comprise investigations of electronic and optical processing methods which can be integrated with imaging sensors to perform low level image enhancement and feature extraction within the sensor. An adaptive CCD processor has been developed for first stage image processing. The smart sensors research work is being performed by the Hughes Aircraft Company Research Laboratories in Malibu, California under subcontract to USC.

Section 2 of this report details the research effort on the image understanding projects. The image processing research activities are described in Section 3, and Section 4 covers the work on the smart sensors projects. Section 5 is a compendium of abstracts of recent Ph.D. dissertations covered under this past six month period and Section 6 is a list of publications by project members during the reporting period.

2. IMAGE UNDERSTANDING PROJECTS

The results of the image understanding projects described below are directed toward the eventual development of processing systems for generalized image analysis. Applications of such systems include photointerpretation, vehicle guidance, visual tracking, cartography, and many other areas of military interest.

Our research study is organized according to the structure of a conceptual image understanding system. This system consists of a feature extraction stage which detects and measures primitive features such as edges and texture regions from an input image. Next the primitive features are clustered into meaningful symbols such as object boundaries, segments of some common attribute, or basic shape structures. Such clustering then is used to completely segment the scene of interest. Finally image symbols are interpreted in terms of their semantic relationships to produce a concise quantitative description of the original image. Some extra-image knowledge base is assumed available for guidance of all elements of the image understanding system.

The philosophic approach to the study has been to bring together a research team skilled in digital signal processing and concepts of artificial intelligence. Although the research effort at USC involves all stages of the overall image understanding system, emphasis has been placed on the feature extraction, clustering, and symbolic description subsystem elements. Rapid progress is being made in these areas to assist the research in the higher stages of the image understanding system.

2.1 EVALUATION OF A SIMPLIFIED HUECKEL OPERATOR

Ramakant Nevatia

Introduction

An edge detector due to Hueckel <1,2> has been in wide use. Though no definitive studies have been made, its performance is considered to be superior to that of other edge detectors (e.g. see <3>, see <4> for a contrary conclusion). However, computation required for a Hueckel edge detector is considerably larger than for other types of operators. The central part of the Hueckel edge detector involves approximation of a two dimensional image intensity pattern by expansion into a truncated Fourier series. Mero and Vassy described a method of simplifying the approximation by using fewer bases functions for expansion <5>. In this report, we examine a generalization of this approach and effect on performance of the reduction in the number of bases functions. The main experimental conclusion is that such reduction in computation is achieved at a significant loss in performance, if the images in use are noisy.

Operation of the Hueckel Edge-Line Detector

The basic process of Hueckel edge operator consists of optimally fitting an ideal edge-line to the image intensity values in a small circular neighborhood. The ideal edge-line is determined by a 7-tuple of parameters, three parameters determining the brightness levels (b_-, t_- and t_+, as shown in figure 1) and the other four parameters determine the position, orientation and the width of the line. The

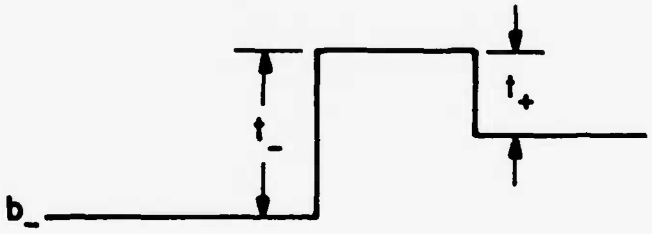


Figure 1. Intensity profile of an ideal edge-line.

fitting process consists of determining the values of the seven parameters for a best fit with the image intensities. Ideally a tuple of parameters is to be computed such that

$$N = \| I - S(\text{tuple}) \| \qquad (1)$$

is a minimum, where I is a vector representing image intensities and S is an ideal edge-line.

This minimization process is approximated ty expansion of both the input image disk and the edge-line in an orthogonal Fourier series. Let a_i be the coefficients of expansicn for the image and s_i be the coefficients for an ideal edge-line. The minimization of (1) is then approximated by choosing a tuple such that

$$N = \sum_{i=0}^{8} (a_i - s_i)^2 \qquad (2)$$

is minimized. Hueckel gives arguments for the optimality of the chosen series <2>. A crucial decision is to use only the first nine terms of expansion. Two reasons are given for the choice of this number:

(a) Higher order terms correspond to noise in the image and should be ignored.

(b) An analytical solution to the minimization problem is found using nine terms.

Decision as to the presence of an edge is based on the amplitude of the computed step, as well as the degree of fit, determined by N in eq.(1) above.

Simplification of the Operator

The computation of Hueckel operator can be reduced substantially if the number of terms used in the expansions above is reduced. An edge-line is completely specified by seven parameters. Further, absolute brightness level (b_-) is of little interest for edge computation; only six Fourier coefficients are needed to compute these parameters. Also, the problem of minimization in eq.(2) above reduces to that of determining six unknowns in six equations (details are in the Appendix).

Only four parameters (orientation, position and step size) are needed if only step edges are of interest. Mero and Vassey describe an approach to determining these parameters for a square region of the image <5>. They conclude that expansion in only two bases functions is needed to determine the orientation of the step. The position and step size are then determined by an approximation.

We have generalized the procedure to determine the four parameters for a step from four coefficients of expansion, used by Hueckel, for a circular neighborhood and determining the six parameters for a step-line from six coefficients of expansion.

In either case, as far as the truncated expansion is concerned, the signal can be fit perfectly with an ideal edge or an ideal edge-line. Thus, no minimization computation is necessary.

The important question is the effect of ignoring some terms in the orthogonal expansion of the signal. It is clear, that if the signal indeed consists of an ideal edge element, the fewer terms of expansion are sufficient to characterize the signal exactly. Results of experimental evaluation signals containing noise are presented next.

Experimental Results

The input for these experiments consists of a picture with two vertical edges; the intensity profile is as illustrated in figure 2. A varying amount of random, Gaussian noise is added to this image. Results are presented for three images, figures 3a, 4a and 5a, with step size to noise variance ratios of: 10, 4 and 2.

The three images were processed by three versions of the edge detector:

1. Original Hueckel edge detector (operator 1)

2. Edge line detector using six coefficients (operator 2)

3. Edge detector using four coefficients (operator 3)

Determining the presence of an edge requires the use of a threshold on the computed step amplitude. Use of a higher threshold will allow fewer noise edges but also lose some of the desired edges. For the results presented, the thresholds were determined for the best subjective performance. The differences in the performance of three operators are so marked that a more careful control of thresholds

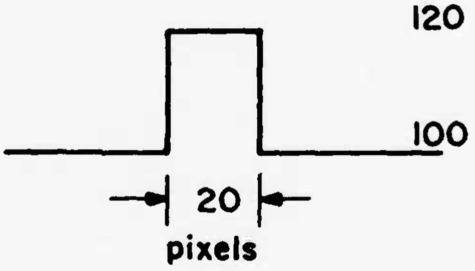
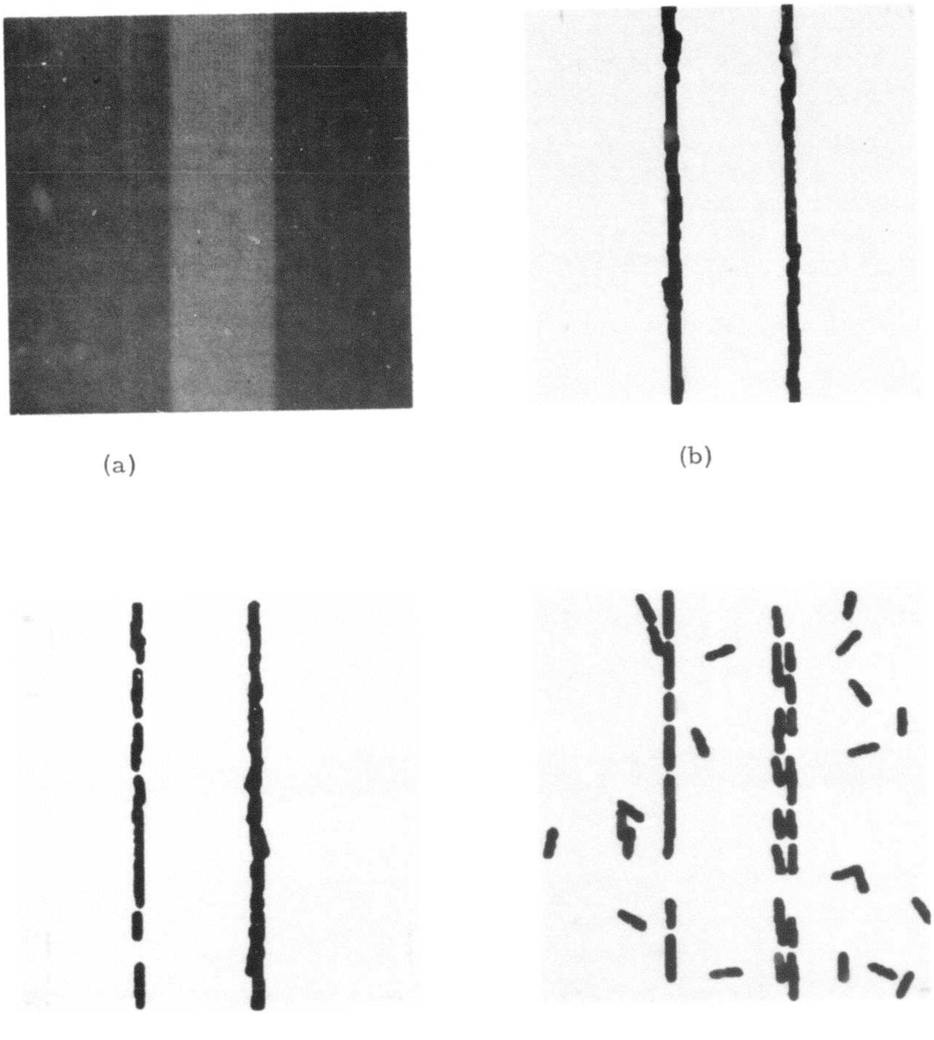


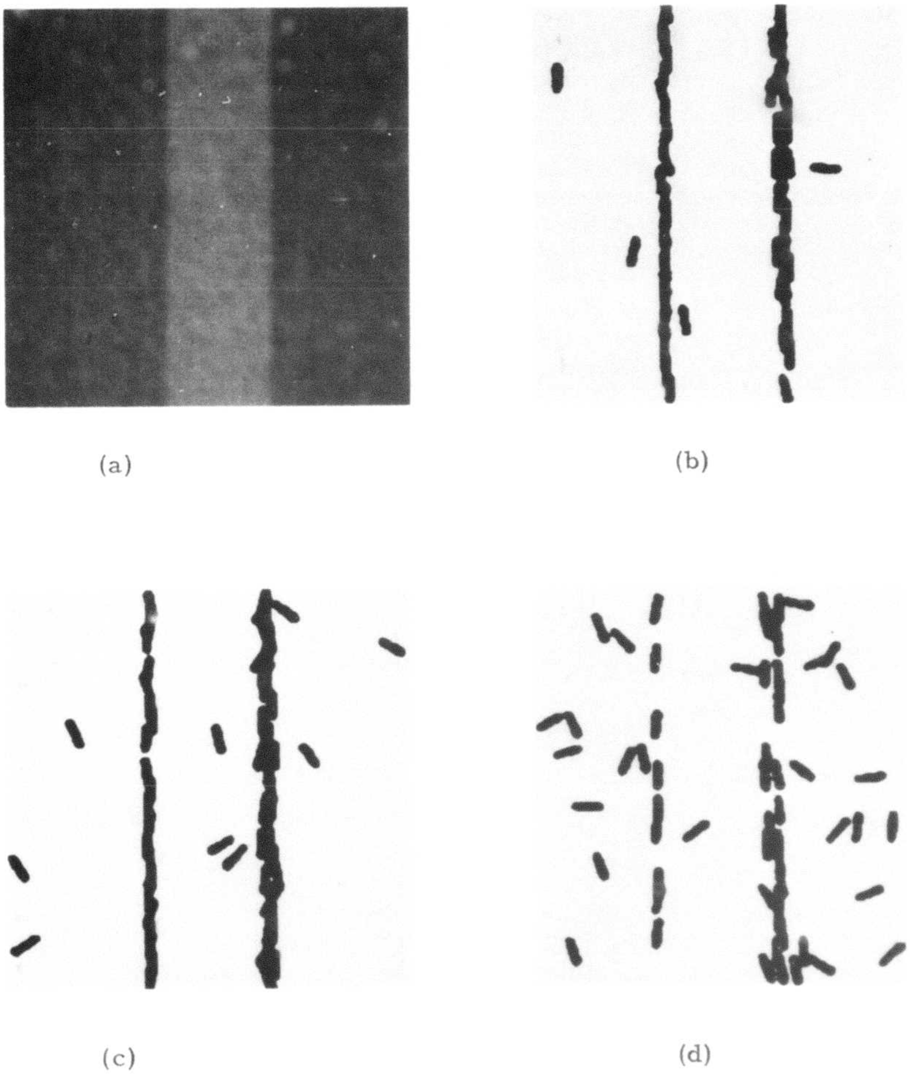
Figure 2. Intensity profile of a test step (without noise)



(a) (b)

(c) (d)

Figure 3. Low noise image and detected edges



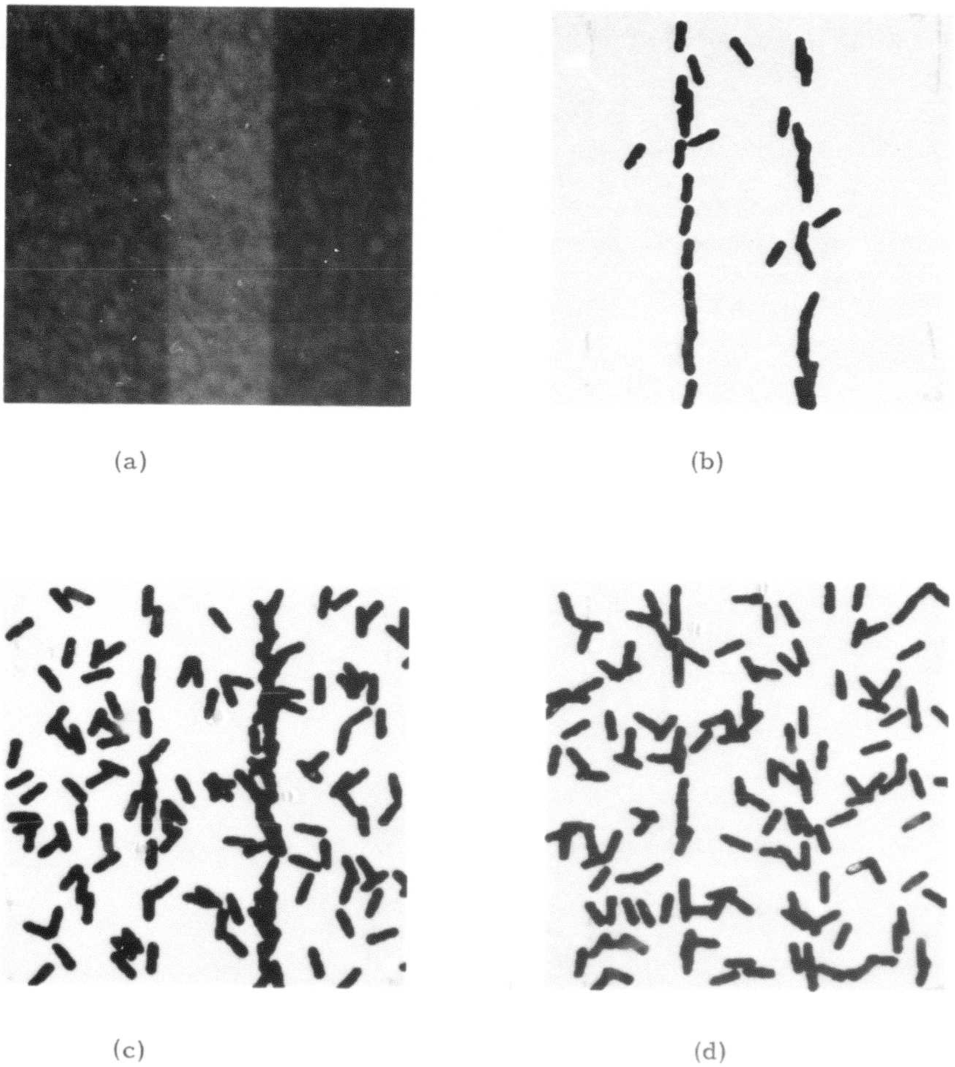
(a)

(b)

(c)

(d)

Figure 4. Medium noise image and detected edges



(a)

(b)

(c)

(d)

Figure 5. High noise image and detected edges

(e.g. by requiring the total number of detected edges to be the same) was not considered necessary.

Figures 3,4 and 5 show the results of the three operators on three images. (The edge detector was applied every three pixels along every third row. Each application looks at a neighborhood approximately circular and eight pixels in diameter. There is considerable overlap of regions in two neighboring applications, resulting in multiple edges for the same step. The phenomenon is pronounced in figures 3, 4 and 5 for the right edge only, because of accidental placements.)

It is clear that the performance of operator 1 declines gradually with the increase in added noise. Operators 2 and 3 on the other hand, perform reasonably well for high signal to noise ratio, but very poorly for images with high noise. Further, the performance of operator 2, using more coefficients of expansion, is superior to that of operator 3. (In absence of any noise, all three operators perform perfectly.)

Similar results are obtained for detection of line edges (i.e. when the width of step is small, say two or three pixels wide).

Conclusions

The results shown above are somewhat surprising. It is tempting to think that as the ideal step is completely determined by four parameters, four bases vectors can be found to span the space of ideal step elements. Unfortunately, the space spanned by the ideal steps is not a complete space (the addition of two arbitrary steps is not another step). (For the special case of edges defined on a 3 x 3 square neighborhood, a complete two dimensional sub-space spanned by the edges has been determined by Frei and Chen <6>).

It is concluded that the simplification of Hueckel edge detector, by ignoring some of the terms of Fourier expansion results in unacceptable loss in performance, if images are likely to be noisy. (Unfortunately, Mero and Vassey present no results on comparable images in their claim that the performance is not compromised.) Further, it raises questions as to the effect of using only nine terms in Hueckel's original operator and also the effects of similar approximations using other series for expansions, (e.g. see <7>).

Appendix: Details of approximations

Familiarity with references <1,2> is assumed here and the notation used therein is used here without much elaboration. The nine coefficients of expansion for an ideal edge-element are as follows:

$$s_0 = (3\pi)^{\frac{1}{2}}(32b_- + t_-(16 - 21r_- + 7r_-^3 - 7r_-^5 + 5r_-^7) + t_+(16 - 21r_+ + 7r_+^3 - 7r_+^5 + 5r_+^7) / 56 \tag{3}$$

$$s_1 = \lambda_- r_- + \lambda_+ r_+ \tag{4}$$

$$s_2 = (\lambda_- + \lambda_+) c_x \tag{5}$$

$$s_3 = (\lambda_- + \lambda_+) c_y \tag{6}$$

$$s_4 = 2^{\frac{1}{2}}(\lambda_- r_- + \lambda_+ r_+)(c_x^2 - c_y^2) \qquad (7)$$

$$s_5 = 8^{\frac{1}{2}}(\lambda_- r_- + \lambda_+ r_+)\, c_x c_y \qquad (8)$$

$$s_6 = 5^{\frac{1}{2}}(\lambda_- r_-^2 + \lambda_+ r_+^2)\, c_x \qquad (9)$$

$$s_7 = 5^{\frac{1}{2}}(\lambda_- r_-^2 + \lambda_+ r_+^2)\, c_y \qquad (10)$$

$$s_8 = \lambda_- r_-(0.5 - 2.5r_-^2) + \lambda_+ r_+(0.5 - 2.5r_+^2) \qquad (11)$$

where r_-, r_+ stand for the positions of brightness transitions, b_-, t and t_+ are as shown in figure 1, $\lambda_{\pm} = (3\pi)^{\frac{1}{2}} t_{\pm} \; (1-r_{\pm}^2)^2/4$, and c_x and c are the direction cosines of the edge orientation (see <1> fo details). Consider two cases separately:

1. Step edges only:

The ideal edge-line can be converted to a step by settin $t_+ = 0$ (hence $\lambda_+ = 0$). We are now interested in determining λ_- r_-, c_x and c_y. We need use only eqs.(4), (5) and (6) and henc need only three coefficients.

2. Step and line edges:

Six parameters need to be determined now and we need si equations. Several subsets of the above nine equations wil suffice. In particular, eqs.(4) through (9) are sufficient. Th solution can proceed by determining c_x and c_y from eqs.(5) an

(6) (note: $c_x^2 + c_y^2 = 1$). Parameters λ_-, λ_+, r_- and r_+ can now be determined from four equations: (4), (5) or (6), (7) or (8), and (9) or (10). These equations can be solved in exactly the same way as in reference <1>, Appendix A. (Note that computation of c_x and c_y using all nine coefficients requires iterative solution of a fourth order polynomial.)

References

1. M.H. Hueckel, "A Local Visual Edge Operator Which Recognizes Edges and Lines," Journal of the ACM, Vol. 20, No. 4, 1973, 634-647.

2. M.H. Hueckel, "An Operator Which Locates Edges in Digitized Pictures," Journal of the ACM, Vol. 18, No. 1, 1971, 113-125.

3. B. Bullock, "The Performance of Edge Operators on Image With Texture," Hughes Research Laboratories Technical Report, Malibu, October 1974 (to be published in Computer Graphics and Image Processing).

4. J.R. Fram and E.S. Deutsch, "On the Evaluation of Edge Detection Schemes and Their Comparison with Human Performance," IEEE Transactions on Computers, C-24, 1975, 616-628.

5. L. Mero and Z. Vassey, "A Simplified and Fast Version of the Hueckel Operator for Finding Optimal Edges in Pictures," Advance Papers of the Fourth International Joint Conference on Artificial Intelligence, 1975, 650-655.

6. W. Frei and C. Chen, "A New Class of Edge and Feature Detection

Operators," in University of Southern California USCIPI Report 660, March 1976, 22-40.

7. K. Paton, "Picture Description Using Legendre Polynomials," Computer Graphics and Image Processing, Vol. 4, No. 1, 1975, 40-54.

2.2 APPENDING OF LINEAR EDGE SEGMENTS WHICH CONTAIN COMMON EDGE POINTS

Lee W. Martin

In an image, there are generally sub-areas which stand out from the background. They may be classifiable as straight lines, curves, boundaries, etc. and they may be given the attributes of a feature, i.e. something that can be used in describing the image scene, using a set of image pixels as a single element. That element, or feature, has a more global information content than the stem of the pixels because of a recognizable interpixel relationship. This report considers the appending of linear segments to create longer, and possibly more identifiable edge segments to be used as features.

Elongated edge segments generally correspond to areas or objects of interest in a scene, e.g. partial boundaries of objects, identifiable segments along an edge (perhaps functions, or particular occurrence of an edge pattern). As such this feature would be useful for stereo analysis (small angle stereo) or scene segmentation. In the latter determine areas of the image where further (second level) processing would be most fruitful, e.g. determine segment continuability criteria determined by the feature segment. In small angle stereo analysis, features in a primary image will generally have a high correspondence with a segment edge which lies on similar points in the complement image, allowing identification of correspondent edge segments. These elongated edge segments satisfy some criteria for being selected as features as previously stated <1>. They have

(a) relative invariance under slightly dissimilar imaging conditions and small perspective changes.

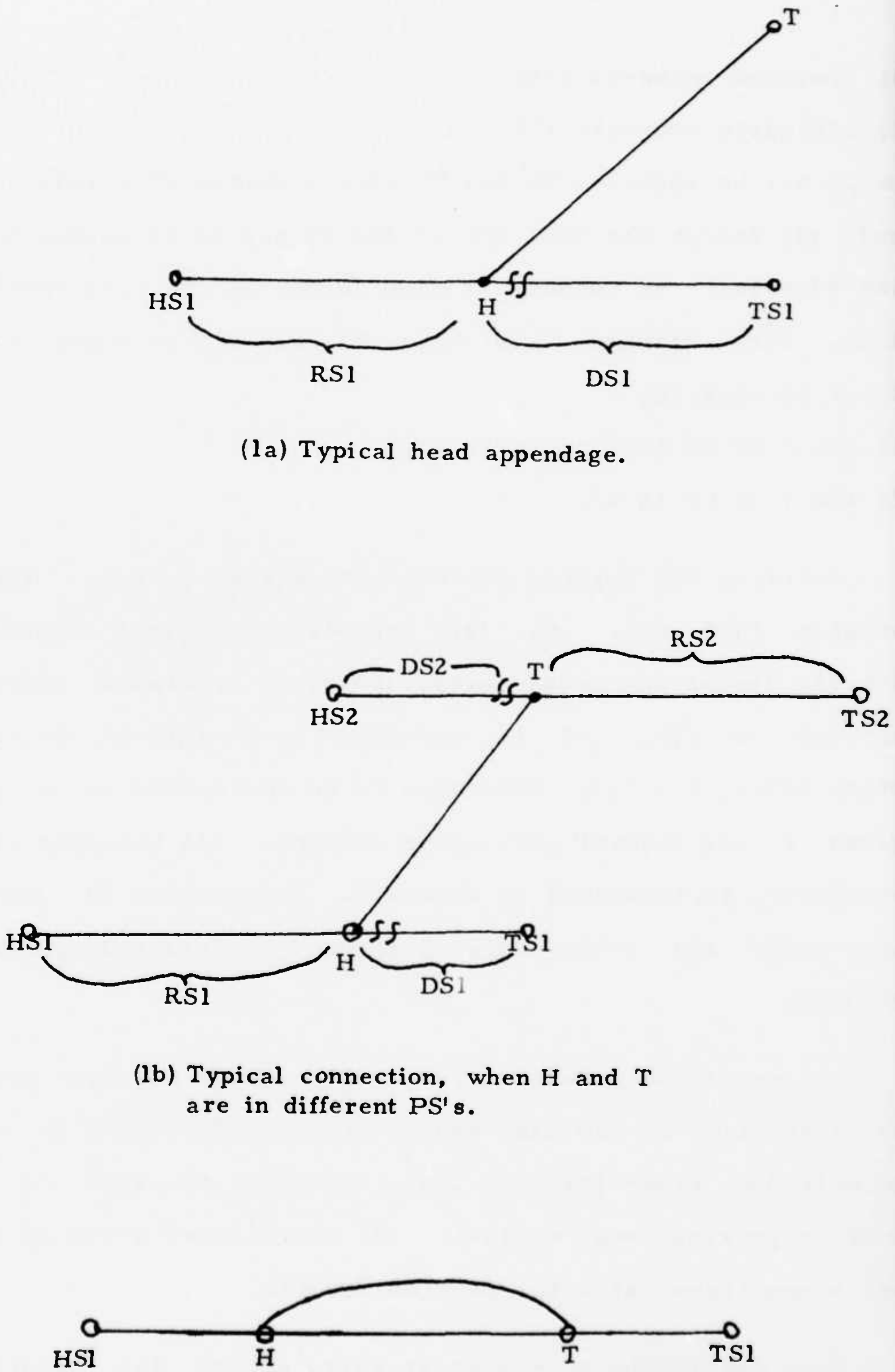
(b) reproducibility of the feature

(c) uniqueness, or at least the set of features which have common attributes should be small, to be useful in complement image matching

(d) correspondence with areas of interest in the image, e.g. object boundaries.

Reference <1> describes the steps leading to the attainment of a linear segment map. Basically a Hueckel <2> edge operator processes the entire image returning an edge point map, which in turn is processed to link edges within orientation and proximity constraints <3>. At this point, edge segments which are continuous in the image are broken into subsegments which lie along pre-selected orientations. Generally, if the edge curve is smooth, edge points near the end of linear segments will also occur in linear segments occurring in the same, or adjacent angles. Thus we have "points of continuity" with which we may link linear edge segments to create more general edge segments of greater length.

The typical links possible are shown in figure 1. Segments are classified as



(1a) Typical head appendage.

(1b) Typical connection, when H and T are in different PS's.

(1c) Typical case where H, T are in same PS. Then CS must be absorbed into S1.

Figure 1. Append Schematics

(1) previous segments (PS)

(2) candidate segments (CS).

The CS may be appended to any PS with a common edge point. Only the head (H) and/or the tail (T) of the CS may be in common with some PS. Each edge point in the set of edge points in the PS's (previously used edges, PUE) belongs to a unique PS, but may re-occur in some future CS. A PS contains

(1) the H of CS designated as S1,

(2) the T of CS as S2.

Notation HS2 implies the head of segment 2, i.e. the PS which contains T(of CS). D, (R), preceding a segment identifier implies possible disconnect subsegments, (possible remainder subsegment i.e. appended to CS). CS is appended to S1 (S2) if the length (CS) > length (DSI), I = 1,2. Otherwise CS is identified as a new segment (given a new segment identifier number). All possible linkages were considered, as indicated in figure 1. The program to accomplish the inter-angle and intra-angle links is complete and approximately 95% debugged.

The preliminary results indicate a slight angular preference in the direction of the step angle (+cc,-c) if linking is not done over the full 360 range (because links can only be made to PS's which occur at previous step angles). All anticipated modes of linking have been accomplished with the desired results.

Upon completion of the programming effort the results will be applied to stereo pairs for evaluation.

References

1. L.W. Martin and R. Nevatia, "Extraction of Prominent Features from Aerial Photographs," University of Southern California, Image Processing Institute Semiannual Technical Report 660, March 31, 1976, pp. 33-100.

2. M.H. Hueckel, "A Local Visual Operator which Recognizes Edges and Lines," Journal ACM, Vol. 20, No. 4, October 1973, pp. 634-647.

3. R. Nevatia, "Object Boundary Determination in Textured Environment," Proceedings of ACM 1975 Conference, Minneapolis, Minnesota, October 1975.

2.3 CONTRIBUTION OF EDGE ORIENTATION OF EDGE DETECTION

Peter Chuan

The purpose of this investigation is to find the class of edge detectors that could find edges with greater resolution. One of the properties that could be looked into is the orientation of edges in the image. The pictures we are most concerned with are pictures of natural images, most of which contain structured objects like trucks and vehicles on a natural background. Figure 1 shows the picture of a tank with a background of marsh. This background, for our purposes of extracting only the tank, is considered noise. However, it should be noticed that this very background which we condemn to be noise actually contains considerable information in itself. Therefore, any good edge operator that will pick up high resclution edges will detect both background and object. The problem facing all edge detectors is to be able to selectively discriminate edge points not of interest. For this, locally adaptive thresholding is considered.

An edge magnitude map and an edge orientation map is produced. Figure 2 shows the edge map thresholded so that the most significant 4000 edge points are displayed. In this example, the Kirsch operator is used. Based on the observation that edge points which do not belong to any structured object should have edge directions distributed randomly, the variance σ^2 of edge directions inside a 3 x 3 window is computed. The edge magnitude map $e(x,y)$ is then modified by the following relation

and the resulting modified edge map $e'(x,y)$ is thresholded to obtain figure 3.

Other information that could be obtained from the edge direction map is the continuity of edge directions. In figure 4, the edge map $e'(x,y)$ is thresholded, where $e'(x,y)$ is the modified version of the edge magnitude map $e(x,y)$ given by the relation

If the edge orientation at point (x,y) is $O(x,y)$, and the orientation of the pixels at the front end and back end of $C(x,y)$ is $O(x_1,y_1)$ and $O(x_2,y_2)$, respectively, then

Figure 1. Tank on a noisy Background

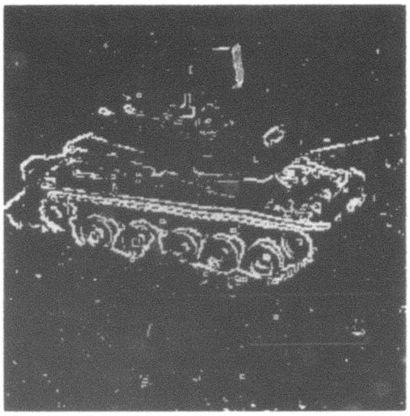


Figure 2. Kirsch thresholded to display 4000 most significant

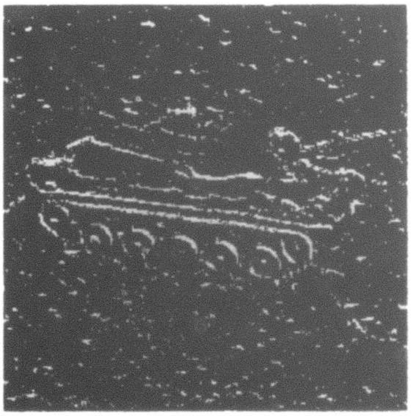


Figure 3. Kirsch with variance modifying function

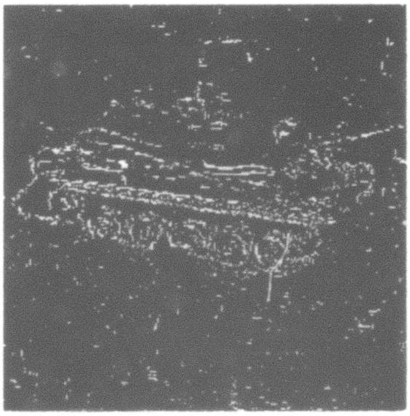


Figure 4. Kirsch with continuity modifying function

Figures 5a, 5b, and 5c show the same three procedures applied to a noisy straight line edge and evaluated with Pratt's figure of merit <1>. The original picture of the line has an S/N ratio of five defined in the convention given by the above figure of merit.

The original Kirsch output has a figure of merit F = 68%. The Kirsch output implemented with directional variance measure has F = 51% and the Kirsch output implemented with continuity modifying function has F = 75%.

From figures 2, 3, and 4 it is obvious that the continuity and variance modifying functions create more backgrcund edges than the original edge magnitude map. However, figure 4 shows that more edges on the side of the tank were detected by the continuity modifying function while the other edges are more or less preserved. Figure 3 shows edges that have less random edge directions and therefore most detected edges appear as part of a chain of edge points. It is not exactly clear as to what extent the orientation information could contribute to detecting fine edges.

References

1. W.K. Pratt, "Figure of Merit for Edge Location," USCIPI Semiannual TR 660, 1976, pp. 85-93.

2. L.S. Davis, "A Survey of Edge Detecticn Techniques," Computer Graphics and Image Processing 4, 1975, pp 248-270.

3. G.S. Robinson, "Detection and Coding of Edges Using Directional

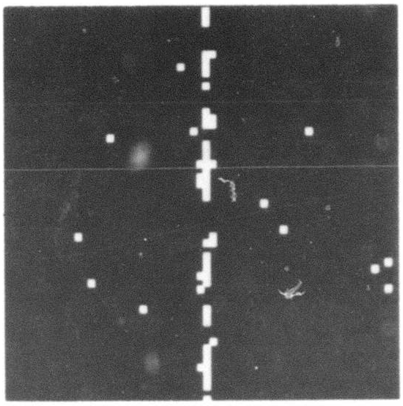


Figure 5a. Kirsch output
F = 68%

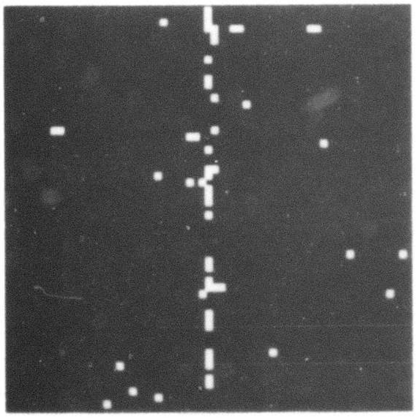


Figure 5b. Kirsch with variance modifying function
F = 0.51%

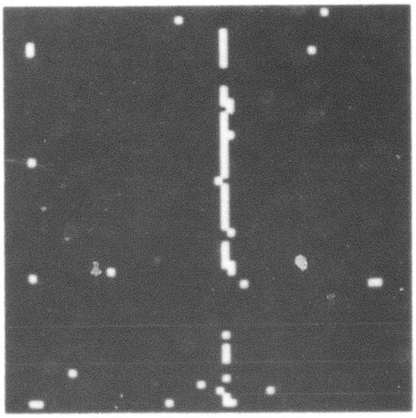


Figure 5c. Kirsch with continuity modifying function
F =75%

Masks," USCIPI Semiannual TR 660, 1976, pp. 41-56.

2.4 SCENE SEGMENTATION BY CLUSTERING

Guy Coleman

The major portion of the effort during the period of this report has been directed at examination of scene features and their relationship to the overall clustering problem. The only intrinsic features of a pixel in a monochrome image are its position (x and y coordinates) and its brightness. All other features are of higher order, that is, they are based upon the relaticnship of the pixel to other pixels in the scene.

A number of different features were computed for overlapping blocks in an aerial photograph (see figure 1). As expected, the larger block sizes virtually destroy the scene detail. These features (see figures 2 through 6 for examples) were computed by averaging the value of the feature over the block.

A different approach was adopted for the tank scene (figure 7) and the armored personnel carrier scene (figure 8). In these scenes, features were computed for several block sizes centered on each pixel. While this approach requires more computation, the intrinsic resolution of the picture is maintained (see figures 9 through 22 for examples). Some of the features (mean, for example) still cause substantial blurring and loss of detail.

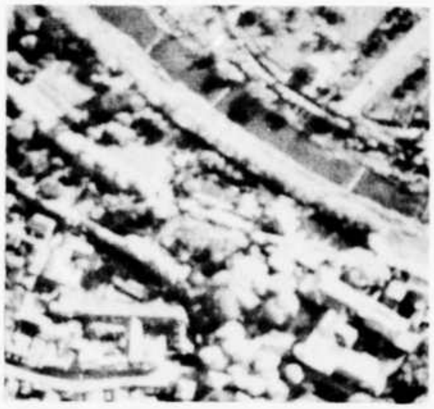


Figure 1. Aerial photograph original

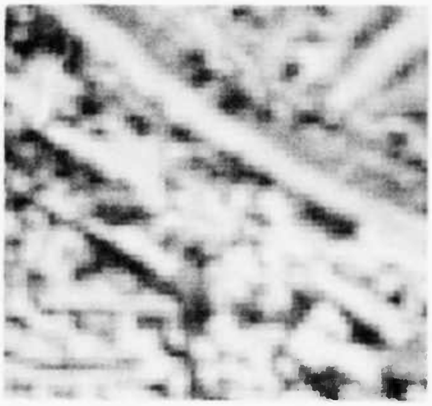


Figure 2. Mean 8 x 8

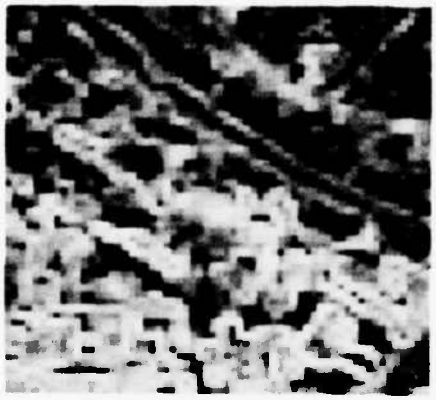


Figure 3. Variance 8 x 8

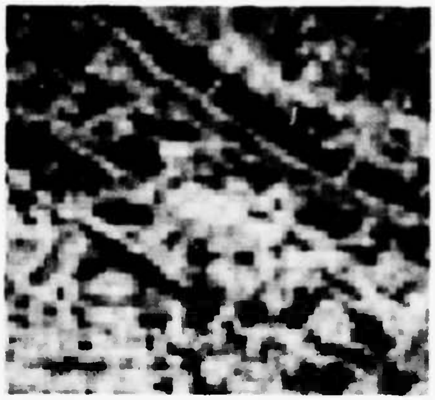


Figure 4. Texture 8 x 8

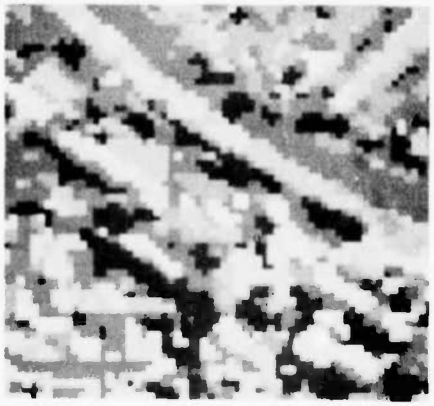


Figure 5. Mode 8 x 8

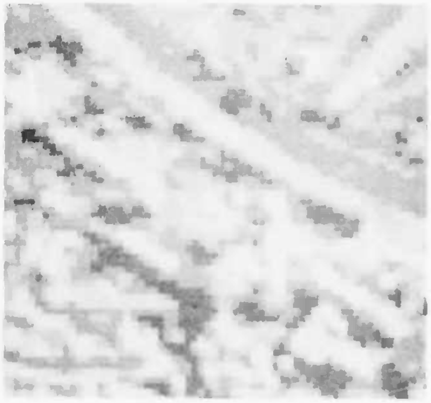


Figure 6. Median 8 x 8

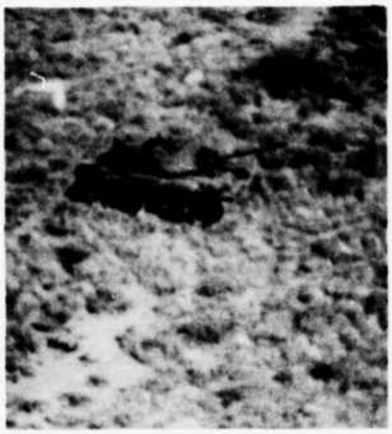


Figure 7. Tank original

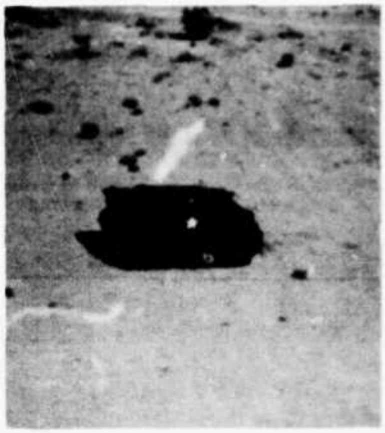


Figure 8. APC original

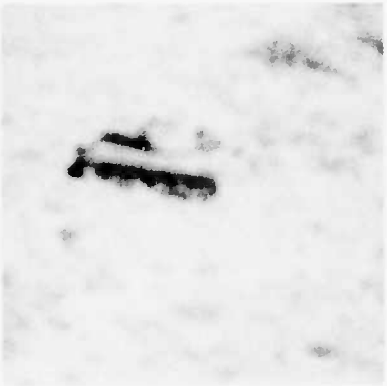


Figure 9. Tank mean 7 x 7

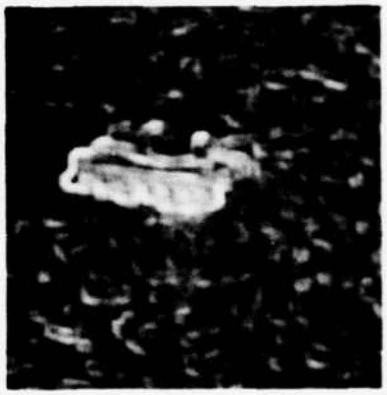


Figure 10. Tank variance

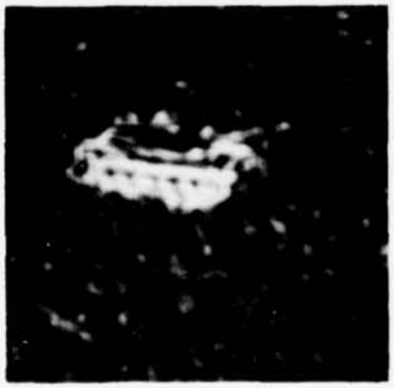


Figure 11. Tank texture 7 x 7

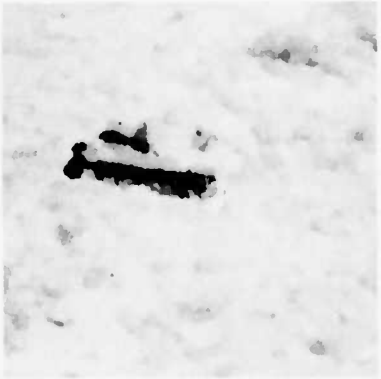


Figure 12. Tank mode 7 x 7

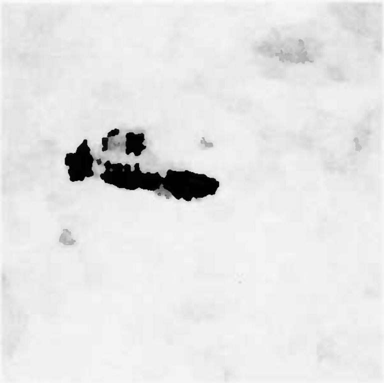


Figure 13. Tank mode 15 x 15

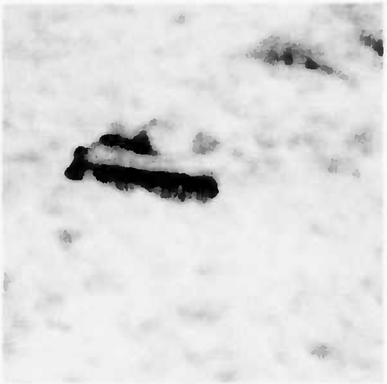


Figure 14. Tank median 7 x 7

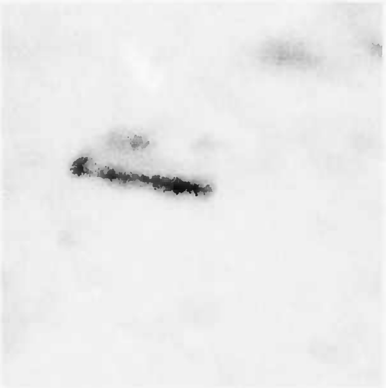


Figure 15. Tank median 15 x 15

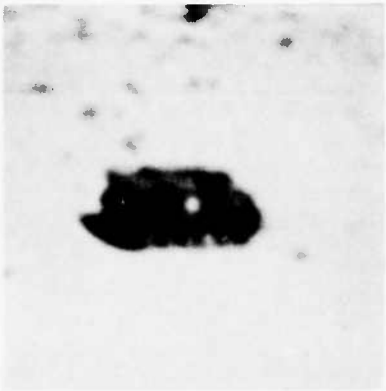


Figure 16. APC mean 7 x 7

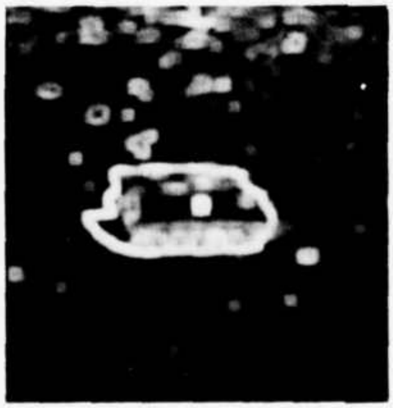


Figure 17. APC Variance 7 x 7

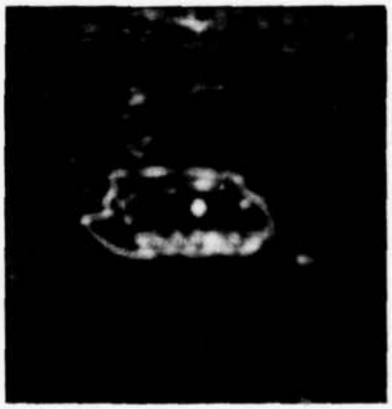


Figure 18. APC Texture 7 x 7

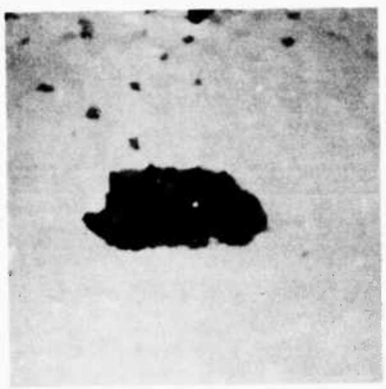


Figure 19. APC mode 7 x 7

Figure 20. APC mode 15 x 15

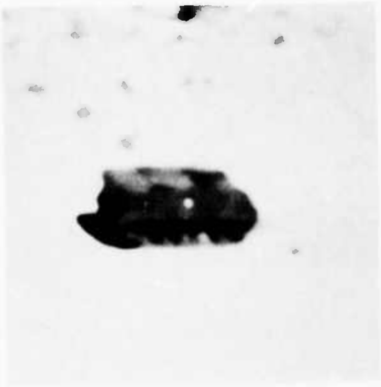


Figure 21. APC median 7 x 7

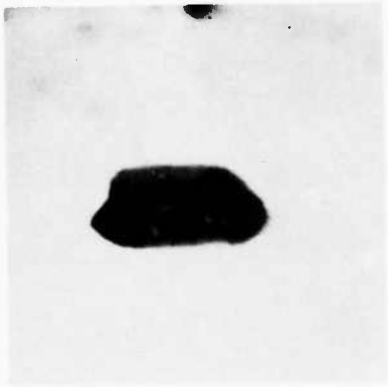


Figure 22. APC median 15 x 15

Some of the features are highly influenced by the value of a f pixels within the block. Examples of this are variance and textu (averaged modified Roberts Cross operator). These features alwa have large values at edge boundaries. Segmentations based on the features would therefore be expected to identify "edge" as a distin region in the segmented image.

The median and mode features are formed by computing the gr scale histogram median and mode in a block centered on each pixel the scene. These features (mode in particular) do not cau substantial degradation of the scene resolution. The median featu does cause some blurring while the mode feature fills in some conca areas and clips some highly convex areas. Nevertheless, featur subjected to modal filtering are expected to substantially improve t classifier performance since they are based on the most frequent cccurring values within the block centered cn each pixel. Textu features having these properties are being investigated.

The activity in progress at present is concerned with featu thinning as a preprocess to clustering. The feature covariance matr was computed and subjected to Karhunen-Loeve transformation. T resultant features are linear combinations cf the original featur and are statistically uncorrelated. The highly correlated set original features will be replaced with an uncorrelated set that substantially smaller. Reduction of the feature space dimension this manner is expected to reduce the amount of computation requir in the clustering algorithm.

The clustering phase of the scene segmentation is presently in a preliminary state. It is anticipated that the overall performance will not be highly dependent on the clustering algorithm chosen and the algorithm chosen will be one having computational efficiency. The procedure under consideration presently will proceed as follows (see figure 23):

1. Initial clustering with a few clusters will be performed using the thinned features computed previously.

2. The "quality" of the clustering will be computed by any of several methods presently under consideration.

3. The number of clusters will be increased (by one perhaps) and steps 1 and 2 repeated.

4. The process will end when the improvement of the cluster quality is less than a threshold.

It has been recognized that many of the linguistic or artificial intelligence methods at higher levels of the image understanding systems rely on preprocessing which produces closed, connected boundaries of the scene segments. The approach described here is based on the vector space model and uses communication theory methods to achieve segmentation. Because of that, a finite mis-classification probability will always exist. It is likely that a transitional procedure will be required to produce closed connected boundaries from the segmented scene before further processing can be performed.

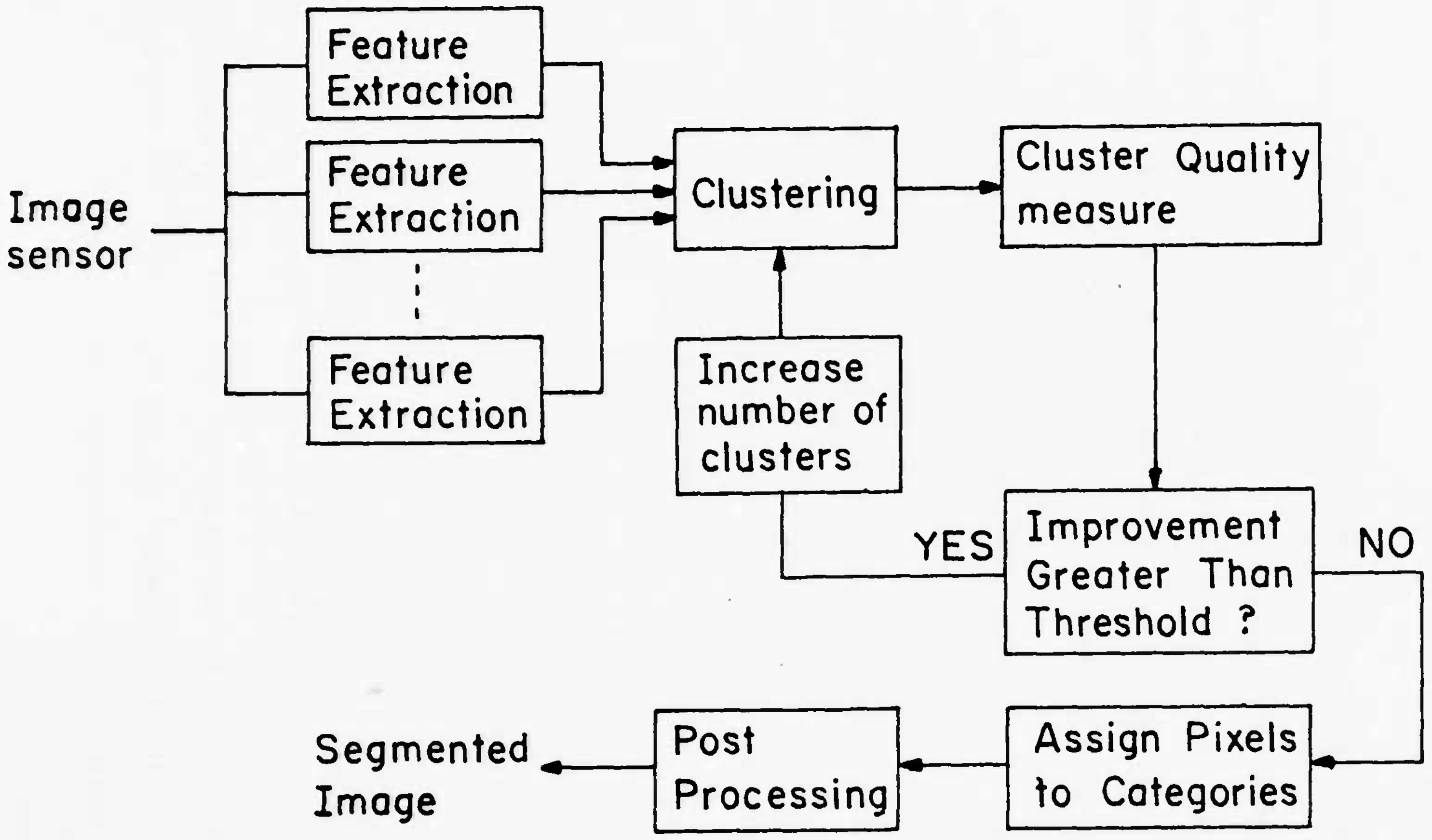


Figure 23. Clustering Procedure

2.5 RECURSIVE ESTIMATICN OF BOUNCARIES

Nasser E. Nahi and Simon Lopez Mora

The purpose of this research is to develop recursive algorithms for the estimation of boundaries of objects in noisy pictures, the available information being limited to the statistics of the object boundary, object and background.

The scanned picture in the absence of ncise is represented by

$$S(k) = \lambda(k)S_o(k) + (1 - \lambda(k))S_b(k) \qquad (1)$$

for $k = 1,\ldots,N^2$ where N is the number of lines in the picture, s_o and s_b denote the intensity values of object and background assumed to be sample functions of two statistically independent, cyclo-stationary random sequences whose first two moments are available, and λ is a binary valued function taking values of 1 or 0 corresponding to points of the image belonging to the object or the background, respectively. The domains of S_o, S_b and λ are the entire picture.

A set of observations

$$y(k) = S(k) + V(k) \qquad (2)$$

is assumed, with s(k) as defined in eq.(1) and v(k) representing a zero mean Gaussian white noise sequence of variance σ^2.

Let us assign to each k two variables w(k) and c(k) representing the width and the geometrical center (as measured from the midpoint on the image) of the object at the associated line. Hence

$$\begin{aligned} w(k+1) &= w(k) \qquad (m-1)N+1 \le k \le m\,N-1 \\ c(k+1) &= c(k) \qquad m = 1, \ldots, N \end{aligned} \tag{3}$$

Since the object width is a positive quantity and the object must be within the image, $w(k)$ and $c(k)$ satisfy

$$\begin{aligned} w(k) &\ge 0 \\ w(k)/2 + |c(k)| &\le N/2 \end{aligned} \tag{4}$$

In terms of the widths and centers, the chosen representation for $\lambda(k)$ is

$$\lambda(k) = \begin{cases} 1 & \text{for} \quad (m-1)N+N/2 + c(k) - w(k)/2 \le k \le \\ 0 & \text{else} \quad (m-1)N+N/2 + c(k) + w(k)/2 \end{cases} \tag{5}$$

The quantities $w(k)$ and $c(k)$ are random sequences with known first and second order statistics. Let

$$\begin{aligned} w(k) &= \tilde{w}(k) + \bar{w} \\ c(k) &= \tilde{c}(k) + \bar{c} \end{aligned} \tag{6}$$

where $\bar{w}$ and $\bar{c}$ are the mean values, respectively. As a first order approximation, the pair $r(k) \overset{\Delta}{=} [\tilde{w}(k)\ \tilde{c}(k)]'$ is assumed to represent a first order Markov sequence,

$$r(k+1) = A_o(k)\, r(k) + B_o(k)\, u_o(k) \tag{7}$$

where $u_o(k) = [u_1\ u_2]'$ is a zero mean white normal sequence with unit covariance, $A_o(k)$ and $B_o(k)$ are matrices cf appropriate dimensions. As a result of eq.(3)

$$\begin{aligned} A_o(k) &= I \qquad \text{(identity matrix)} \\ B_o(k) &= 0 \end{aligned} \tag{8}$$

for all k except $k = mN$, $m = 1, \ldots, N$.

By augmenting eq.(7) with dynamic equations representing the statistics of the object and background it is possible to state a nonlinear MMS estimation problem that can be solved by the method introduced in <1>.

To illustrate the estimator performance a binary picture with an ellipse of unit luminance against a background of zero luminance has been used. The original observation has been corrupted by adding white Gaussian noise of standard deviation 0.5 ($\sigma = 0.5$ or $\rho = 1.0$) in the high signal-to-noise ratio case and noise of variance 4 ($\sigma = 2.0$ or $\rho = 0.25$) for the highly noisy case. Two different estimators were used: I) estimation of the boundaries when $s_o(k) = 1$ and $s_b(k) = 0$ in eq.(1), this condition singles out the boundary estimation itself, see figures 1b, 1c, 2b, and 2c. II) estimation of the boundary as well as the object intensity when the last one is assumed constant and only a priori information on its variance and mean is known, see figures 1d, 1e, 2d, and 2e. In either case a refinement of the estimator consisting of a detection test at the end of every line has been performed. As a result of this detection step, the estimate can be rejected in which case the estimator is initialized or accepted without change. The estimator detector form is referred to as the refined estimator.

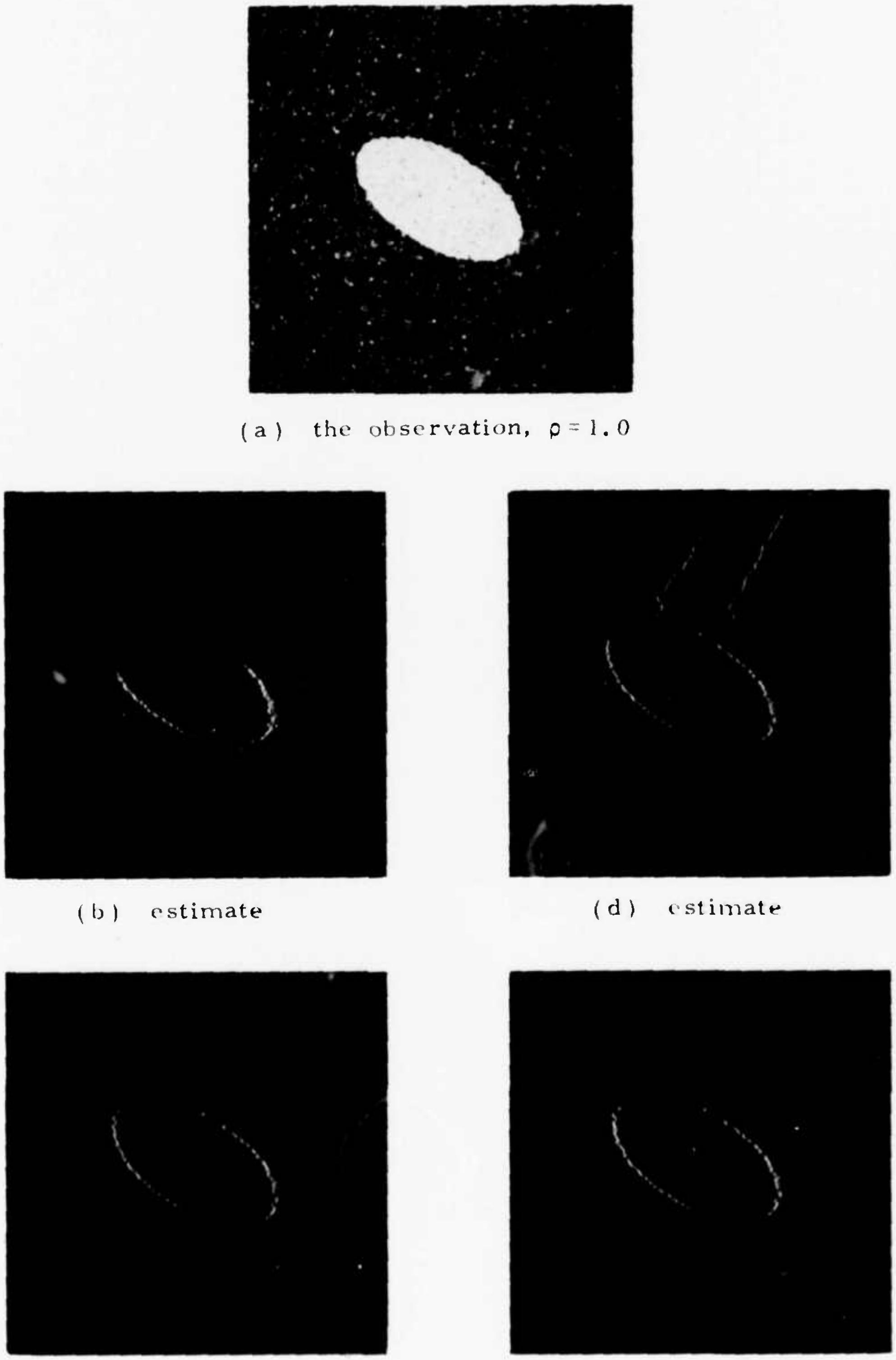
The parameters chosen for these estimators were:

$$N = 256 \qquad \overline{W} = 57 \qquad \overline{C} = 0$$

Case I

$$A_o(mN) = \begin{pmatrix} .926 & 0 \\ 0 & .97 \end{pmatrix} \qquad B_o(mN) = \begin{pmatrix} 7.78 & 0 \\ 0 & 2.91 \end{pmatrix}$$

Case II



(a) the observation, $\rho = 1.0$

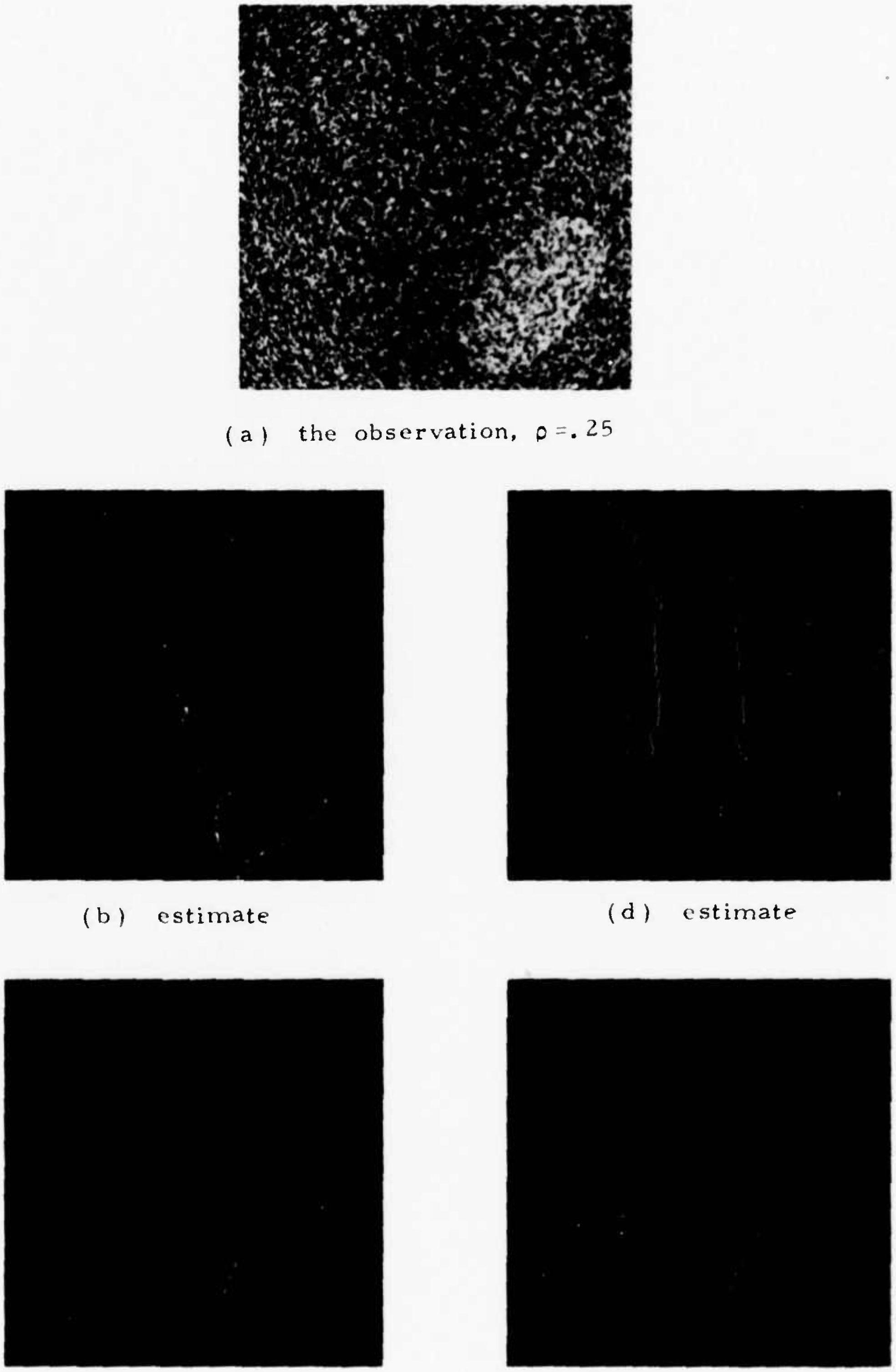
(b) estimate

(d) estimate

(c) refined estimate

(e) refined estimate

Fig. 1. Estimated and original boundaries for known (b, c) and unknown (d, e) object luminance. ρ indicates signal-to-noise ratio.



(a) the observation, $\rho = .25$

(b) estimate

(d) estimate

(c) refined estimate

(e) refined estimate

Fig. 2. Estimated and original boundaries for known (b, c) and unknown (d, e) object luminance. ρ indicates signal-to-noise ratio.

$$A(mN) = \begin{pmatrix} .926 & 0 & 0 \\ 0 & .97 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B(mN) = \begin{pmatrix} 7.78 & 0 & 0 \\ 0 & 2.91 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E\, s_o = 0.5 \qquad E(s_o - E\, s_o)^2 = 0.C5$$

Although the estimator procedure is applicable to non-binary pictures, the way to reduce the possible increase in computation is still under investigation.

Reference

1. N. Nahi and M. Naraghi, "A General Image Estimation Algorithm Applicable to Multiplicative and Non-Gaussian Noise," 18th Midwest Symposium on Circuits and Systems, Montreal, Canada, August 11-12, 1975.

2.6 SHAPE ANALYSIS AND DESCRIPTION

Erica M. Rounds

This research is directed toward constructing region representations which are useful at the interpretation level. A good representation should be largely invariant with respect to location, orientation, and size to facilitate matching with object descriptions in the knowledge base <3>. The process of extracting global features from region descriptions is called here "shape analysis."

Shape analysis can be considered as an intermediate level between the segmentor and the interpreter. Each of these levels deals with image descriptions on a conceptually different basis. For example, regions of homogeneous properties extracted by the segmentor are considered areas by the shape analyzer having geometrical characteristics such as size, perimeter, and shape. On the interpratation level these may correspond to object surfaces. Likewise, global discontinuities in the contour may be associated with vertices of a three-dimensional object.

The shape analyzer operates on the region description produced by the segmentation level. Typically, this consists of the set of points which constitutes the boundary of a region. More precisely, we define the closed boundary of a simple region (nc holes) to be a set of N ordered grid points $\{(x_k,y_k)\}$ such that (x_{k-1},y_{k-1}) is adjacent to (x_k,y_k), $2 \leq k \leq N$, and $(x_N,y_N) = (x_1,y_1)$. Geometrically, the sequence of points corresponds to a clockwise or counter-clockwise traversal of the contour. Two points are said to be adjacent if they are one grid cell apart in a horizontal, vertical, or diagonal direction. Figure 1 shows the eight possible neighbors of a central reference point together with their direction codes <2>.

The region representation will be obtained by a sequence of operations performed on the boundary data such as

(1) gap filling,

(2) local smoothing,

(3) finding global extreme points,

(4) fitting linear or curvilinear line segments between extreme points,

(5) computing the centroid, area and perimeter of the region,

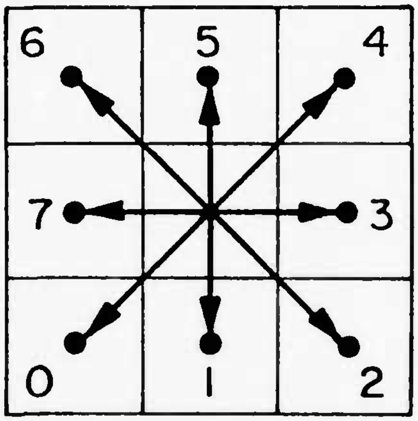


Figure 1. 8 - neighborhood.

(6) computing the convex hull, and

(7) assigning the region to one of a set of standard shapes.

Operations (1) and (2) deal with local features of boundary segments. Gaps can be filled by computing the bias line segment between the two endpoints of a gap and digitizing the intermediate points on the line according to eight-connectedness. An operator for smoothing local extrema is presented in the next section of this report. Operations (3) and (4) seek to divide the closed boundary into meaningful segments. Just as intensity discontinuities are important in edge finding, so curvature discontinuities are significant in object contours <4>. Fitting of curvilinear lines achieves information compression as well as facilitates manipulation at later stages. Operation (6) is useful in identifying concavities in the figure. These would be due to occluding objects. If the fitted lines of the contour are described relative to the centroid and the principal axes <1>, then the representation will be location and rotation invariant. Size invariance can be achieved by normalizing contour segments with respect to area. Invariant properties will reduce the problem of assigning a standard shape to the region. Current research is devoted to (4), (6), and (7).

The operations described above are not only useful for processing boundary points obtained from the segmentor but also for data acquired by manual segmentation. We have developed a number of programs to hand-segment regions interactively from a displayed digital image. The purpose of this work is to construct object descriptions for the knowledge base. Given the resolution of images it is often extremely difficult to manually extract straight continuous boundary segments.

Later portions of this report present an algorithm for generating figures or regions from their boundary coordinates.

References

1. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley & Sons, New York, 1973.

2. H. Freeman, "Analysis of Line Drawings," Proceedings NATO Advanced Study Institute, France, June 1976.

3. E.M. Rounds, "Interpretation and World Knowledge," USCIPI Report 660, Spring 1976, pp. 105-114.

4. K.J. Turner, "Computer Perception of Curved Objects Using a Television Camera," Ph.D. Thesis, University of Edinburgh, 1974.

2.7 LOCAL SMOOTHING OF DIGITAL CONTOURS

Erica M. Rounds

In extracting shape features from region boundaries, a large amount of information is contained in the global discontinuities of curvature, i.e., at the extreme points where a significant change in direction occurs. It is frequently desirable to perform a prior smoothing operation on the boundary points to reduce the number of small local extrema. These may be due to noise or to the particular algorithms used at lower levels.

A very simple technique for smoothing is "Hysteresis smoothing" described by Mason and Clemens <1>. In this approach, x and y coordinates of the boundary points are treated separately. The smoothed curve essentially follows the digitized curve within a threshold of 2A, where A is a given constant. This has the effect of deleting local extrema of height $\le 2A$. Let y_{k+1}be the y-coordinate of the (k+1)st data point, and let $\bar{y}_k$ be its smoothed predecessor. Then $\bar{y}_{k+1}$is determined as follows:

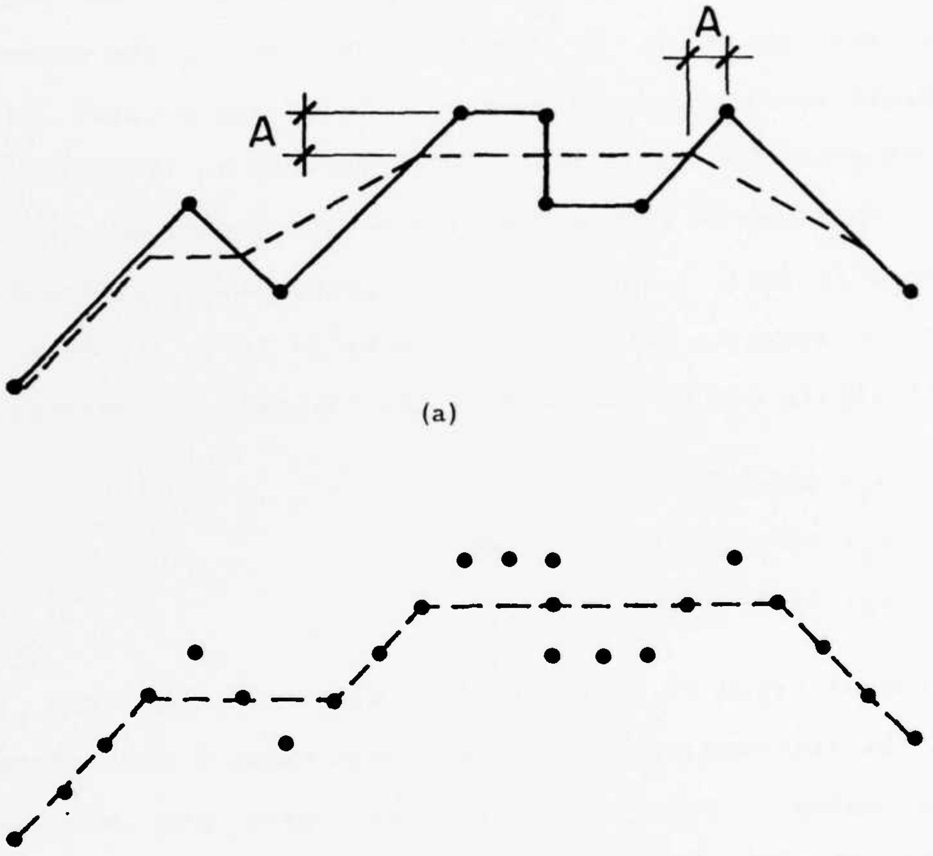
(i) If $y_{k+1} - A \le \bar{y}_k \le y_{k+1} + A$ then $\bar{y}_{k+1} = \bar{y}_k$,

(ii) If $y_{k+1} - A > \bar{y}_k$ then $\bar{y}_{k+1} = \bar{y}_k - A$,

(iii) If $\bar{y}_k > y_{k+1} + A$ then $\bar{y}_{k+1} = \bar{y}_k + A$.

Figures 1a and 1b show the application of Hysteresis smoothing to a curve segment. The dashed line indicates the smoothed curve. Note that this scheme is sensitive to the spacing of data points. In figure 1b where the data points are one grid cell apart (of dimension A), the smoothed curve follows the original one more faithfully. Thus, hysteresis smoothing is sensitive to gaps in digitized curves. It is also highly dependent on the starting point. Since the smoothing process is influenced by the orientation of the boundary segments, it will distort symmetry inherent in the figure. Furthermore, for closed curves the smoothed version may not be closed. (For an example see figure 7a.)

Figure 2 shows a different type of smoothing operation called



(a)

(b)

Figure 1. Hysteresis smoothing

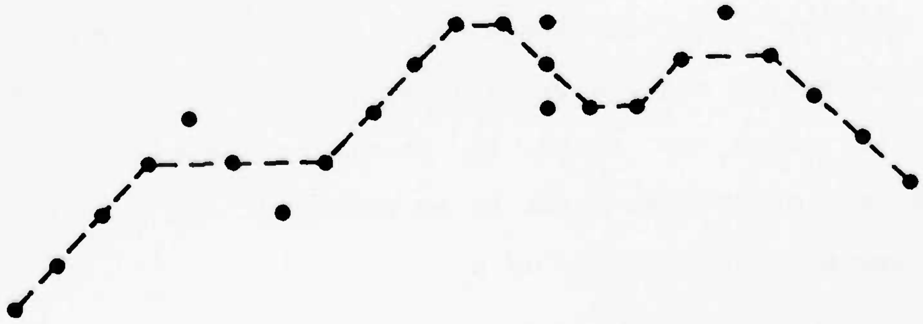


Figure 2. P-S pair smoothing

here "P-S pair smoothing," because it is based on the predecessor and successor points of the data point to be smoothed. We consider all possible P-S pairs in a 3 x 3 neighborhood of the reference point (using eight-neighborhood adjacency). There are a total of 64 such pairs, of which the eight pairs which reverse on themselves have been ignored. That is, we assume that the boundary encloses a region at least one cell wide. The remaining 56 pairs have been classified into six types as shown in figure 3. These types form a basic set from which all others can be obtained by the following operations:

e_1: rotation by 90^o,
e_2: rotation by 45^o, and
e_3: reflection.

Computation of types is very efficient since for each type (excepting I and V) the distance $\alpha(P,S)$ between predecessor P and successor S has a unique value. Thus, at most two tests are necessary. The properties of the P-S types are summarized in table 1 of figure 4. Constant A is the dimension of the grid cell.

The objective of the smoothing operation is to eliminate local sharp corners such as types II, IV, and V. For II and IV, the operation deletes the central reference point. In case of type V, the reference point is moved to create a type I P-S pair. Let $p_i = (x_i, y_i)$ denote the point to be smoothed, and let $\bar{p}_{j-1} = (\bar{x}_{j-1}, \bar{y}_{j-1})$ be the smoothed predecessor of p_i.

Algorithm

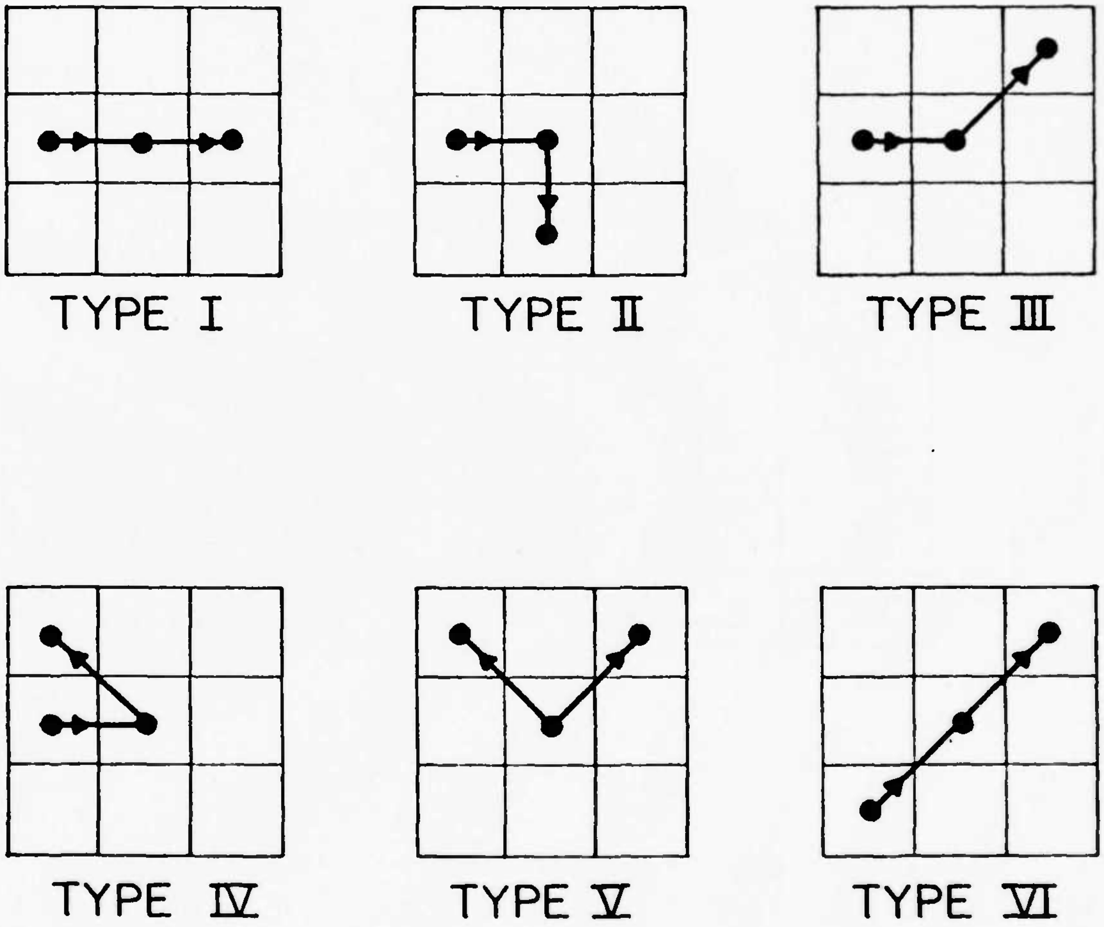


Figure 3. Distinct P-S pairs.

TYPE	d(P, S)	e_1	e_2	e_3	TOTAL
I	$2A$	X			4
II	$\sqrt{2}A$	X		X	8
III	$\sqrt{5}A$		X	X	16
IV	A		X	X	16
V	$2A$	X		X	8
VI	$2\sqrt{2}A$	X			4

Table 1

Figure 4. P, S Type Characteristics

(0) Initialize: Scan the data points until a point p_i is found such that p_i has type I, III, or IV. This is the first smoothed point $\bar{p}_{j-1}$. Set i to i+1.

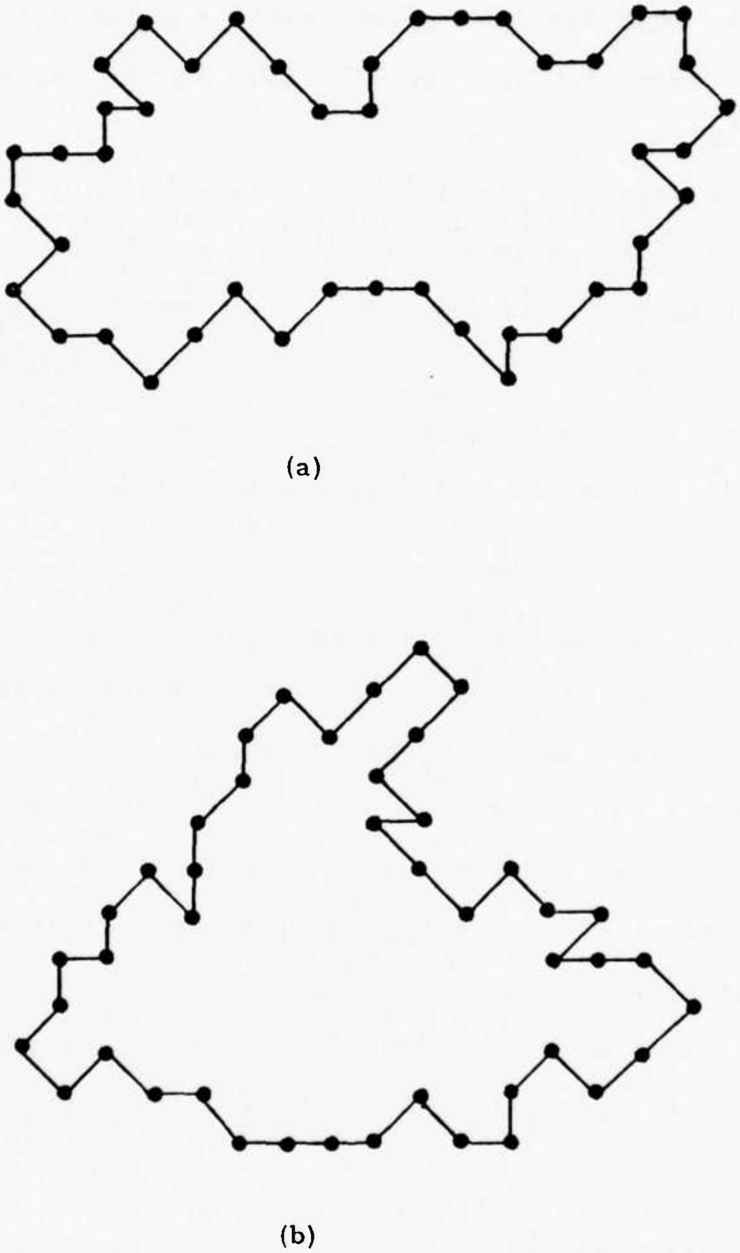
(1) Compute $\alpha = \alpha(\bar{p}_{j-1}, p_{i+1})$ and assign a type to p_i.

(2) If type = I, III, or IV then $\bar{p}_i = p_i$. Go to (3). If type = II or IV then set i to i+1 and go to (1). (This checks the type of the new reference point before assigning the smoothed point.) If type = V then compute $\bar{p}_j$ such that $\bar{p}_j$ has type I. Go to (3).

(3) Set i to i+1, j to j+1 and continue with remaining data points at (1).

The performance of the algorithm is demonstrated on two digitized contours shown in figures 5a and 5b. Figures 6 and 7 compare the effects of Hysteresis smoothing with P-S pair smoothing. In general, P-S pair smoothing tends to produce a curve which has a close resemblance to the original one. Since this method is not direction-dependent it will preserve symmetry. It is also insensitive to gaps in the boundary, since successive points separated by more than 2A are not changed. The choice of starting point (marked in figures 6 and 7 by X) assures that the algorithm will produce a closed contour.

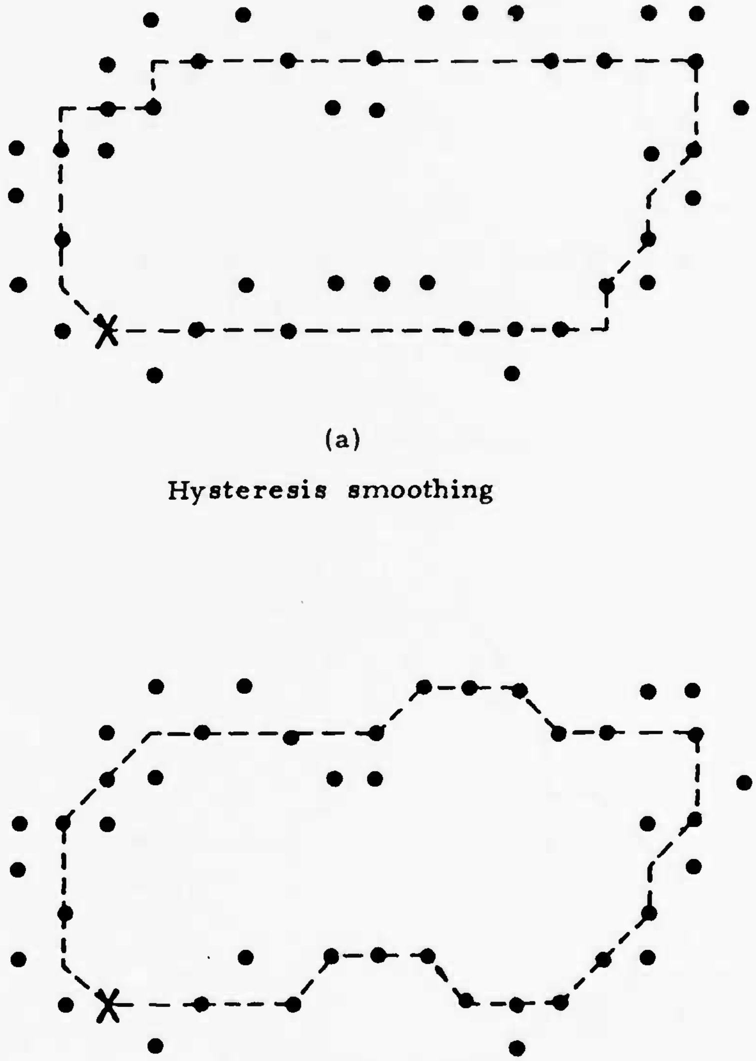
This smoothing algorithm can serve as a preprocessor before applying more sophisticated techniques for curve fitting. Although the 3 x 3 neighborhood is usually not large enough to assess the global significance of some corners, it is felt that the economy of the algorithm outweighs these disadvantages. The large number of



(a)

(b)

Figure 5. Original data points of contour



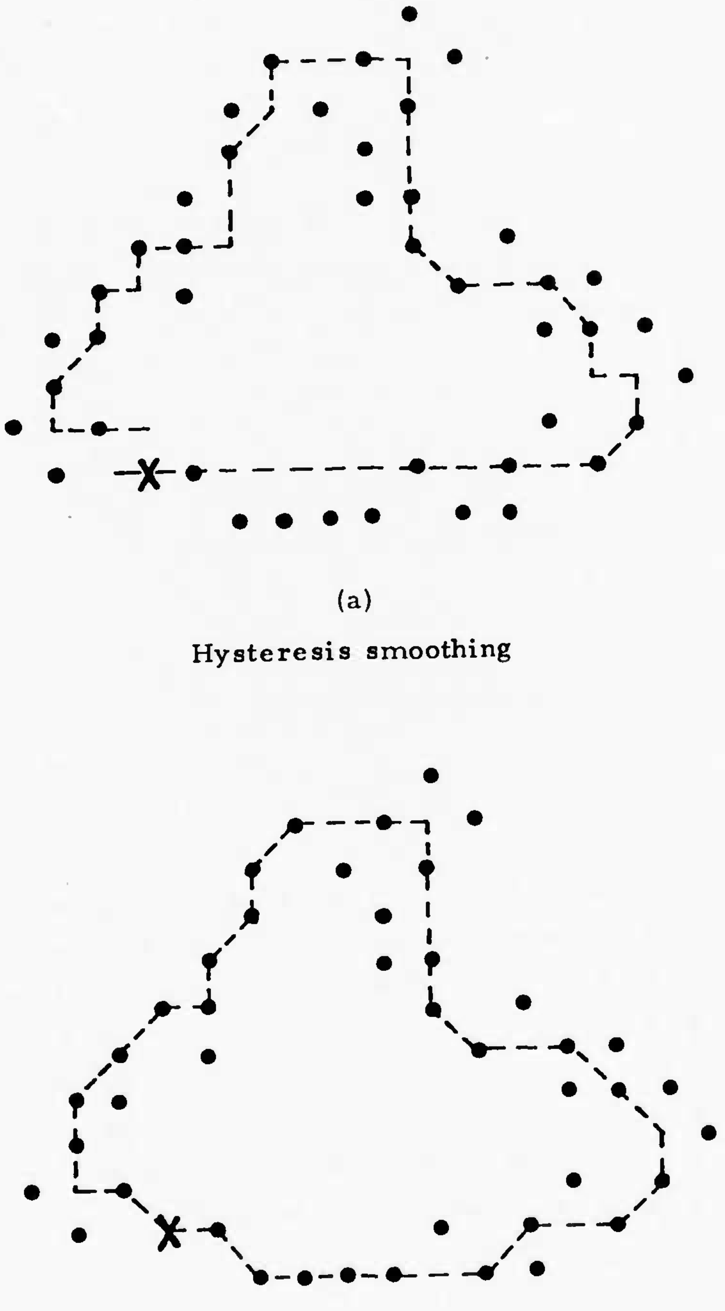
(a)

Hysteresis smoothing

(b)

P-S pair smoothing

Figure 6. Smoothed points of Figure 5. a.



(a)

Hysteresis smoothing

(b)

S-P pair smoothing

Figure 7. Smoothed points of Figure 5.b.

possible configurations which must be considered in a 5 x 5 or 7 x neighborhood would add considerable computational complexity withou producing much better results.

Reference

1. S.J. Mason and J.K. Clemens, "Character Recognition in a Experimental Reading Machine for the Blind," in Recognizing Patterns P.A. Kolers and M. Eden, eds., M.I.T. Press, Cambridge Massachusetts, 1968.

2.8 FIGURE CONSTRUCTION FROM ITS CONTOURS

Erica M. Rounds

A method is presented for constructing a planar, digitized figur from its closed boundary. This can be useful for extracting a desire region from a given scene once its boundary is determined. Anothe potential application is image synthesis where the image i constructed from various regions to which different intensity, colo or texture values have been assigned. Image synthesis could be valuable tool for comparing the processed picture information with th original input scene.

In the following we assume a digitized contour usi eight-neighborhood adjacency and counter-clockwise traversal. T algorithm permits any number of "holes" in the interior of the figu or region. Hole boundary lists are simply appended to the regi boundary list. Figure 1 shows the boundary points for a region with two holes.

The algorithm consists of two steps. During the first step, each boundary point (x_k, y_k) is assigned a type which is stored in a corresponding element (i,j) of a picture matrix P having dimension $(y_{max} - y_{min} + 1) \times (x_{max} - x_{min} + 1)$. The type for (x_k, y_k) is based on the "shape" of the line segment connecting points (x_{k-1}, y_{k-1}), (x_k, y_k), (x_{k+1}, y_{k+1}). Classification into types is straight-forward since each type has a characteristic pattern. This is given by a vector of length 3 whose elements represent the number of neighbors above, on the same horizontal line, and below a given reference point in a 3 x 3 neighborhood. Figure 2 gives the pattern and typical line segments for each of the six types. Figure 3 shows the assignment of boundary types to the points in figure 1. Types 1 and 3 are local peaks (concave or convex), and types 2 and 4 are end segments of a horizontal line. The latter always occur in pairs, possibly separated by type 6 points.

Step 2 is described using figure 4. The interior of the region is indicated by hatching. Each row of P is scanned left to right and interior cells of the figure are filled in, i.e., corresponding elements of P are assigned a specified value. To determine the left and right end or a row of interior points, boundary points are examined as follows:

(1) If the boundary point is approached from the exterior of the region, then

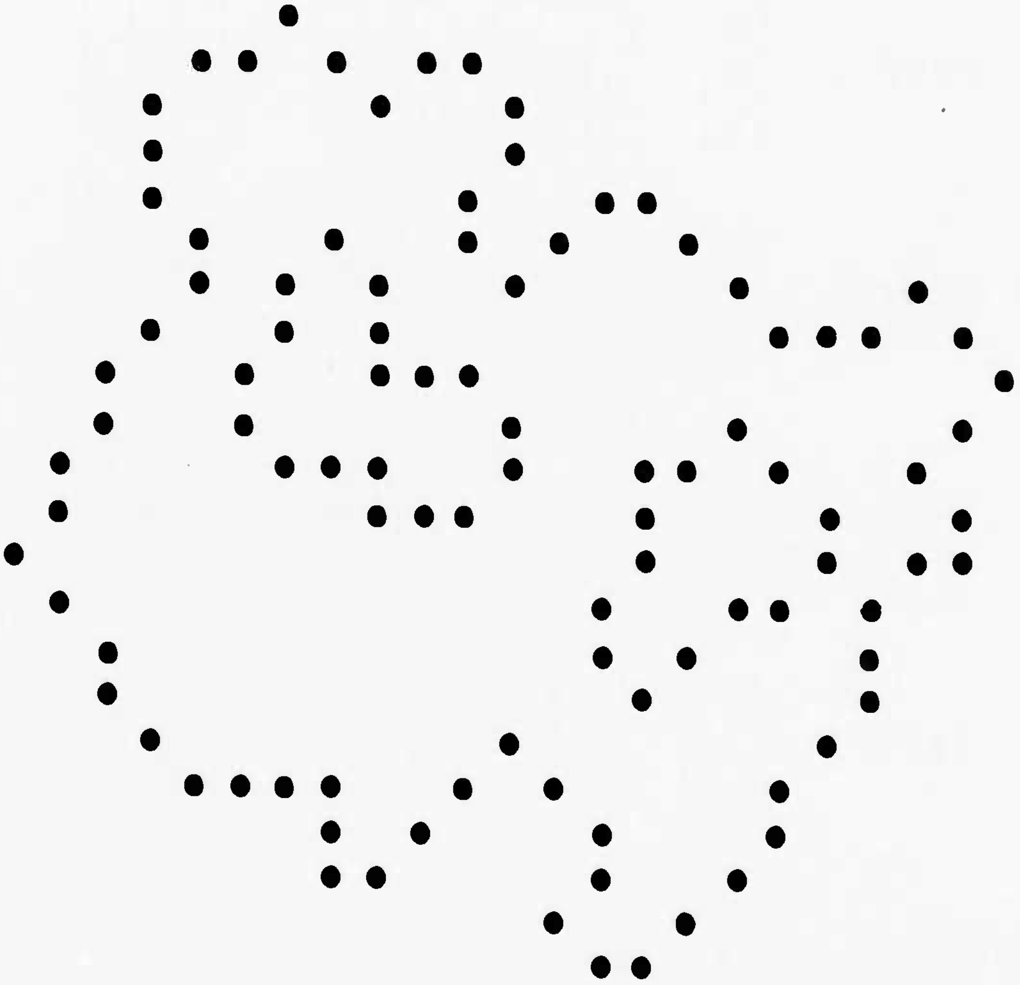


Figure 1. Contour points of regions with holes.

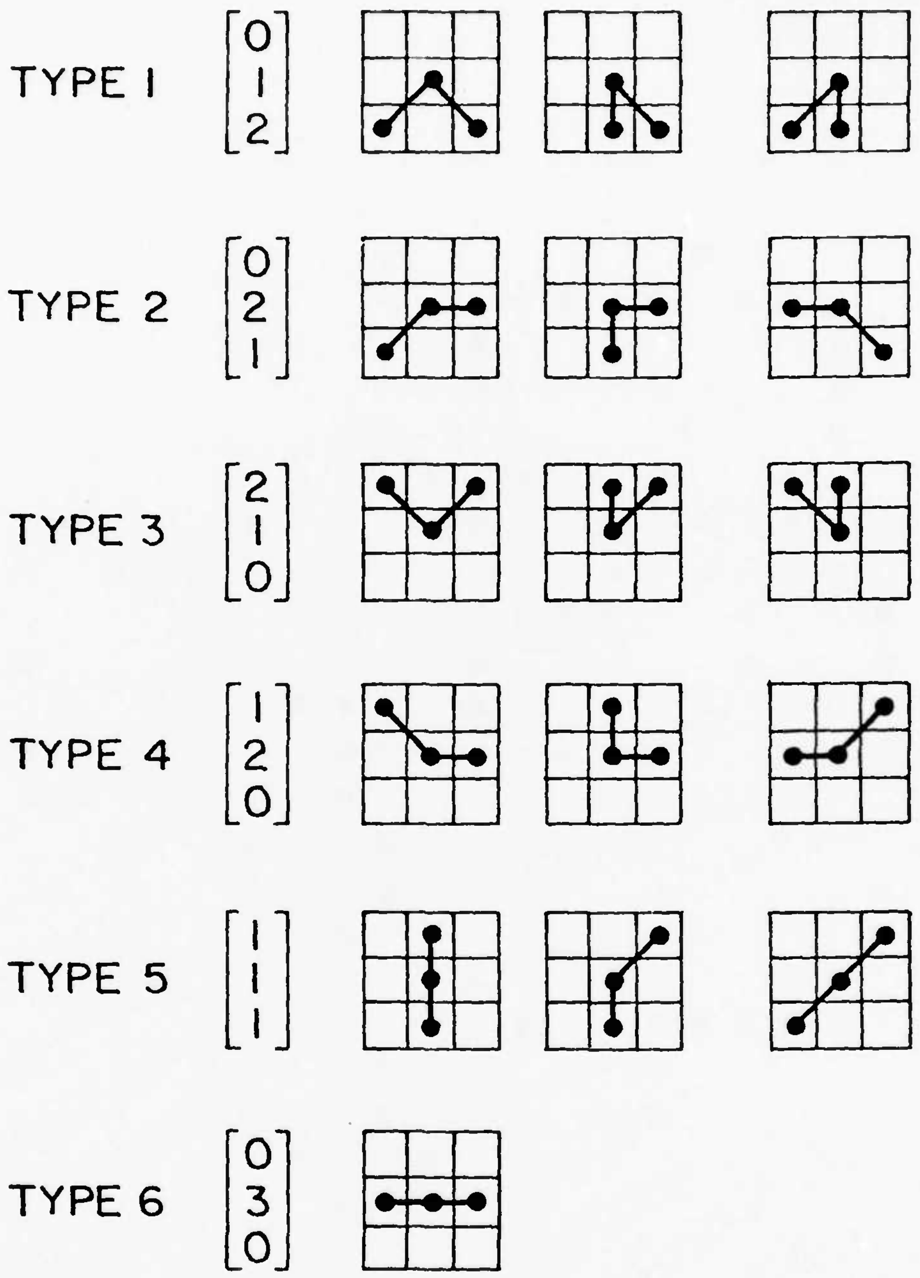

TYPE 1
$\begin{bmatrix}0\\1\\2\end{bmatrix}$
TYPE 2
$\begin{bmatrix}0\\2\\1\end{bmatrix}$
TYPE 3
$\begin{bmatrix}2\\1\\0\end{bmatrix}$
TYPE 4
$\begin{bmatrix}1\\2\\0\end{bmatrix}$
TYPE 5
$\begin{bmatrix}1\\1\\1\end{bmatrix}$
TYPE 6
$\begin{bmatrix}0\\3\\0\end{bmatrix}$

Figure 2. Classification of boundary types.

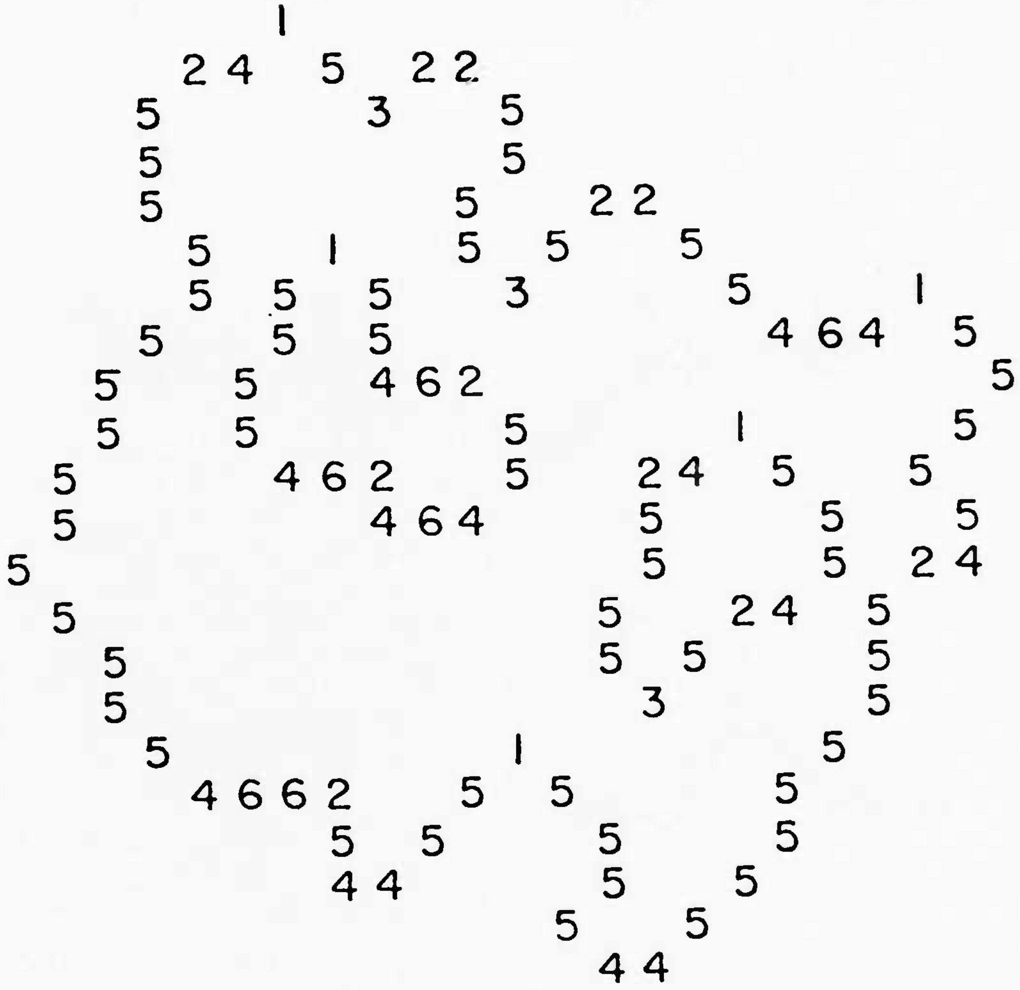


Figure 3. Assignment of boundary types.

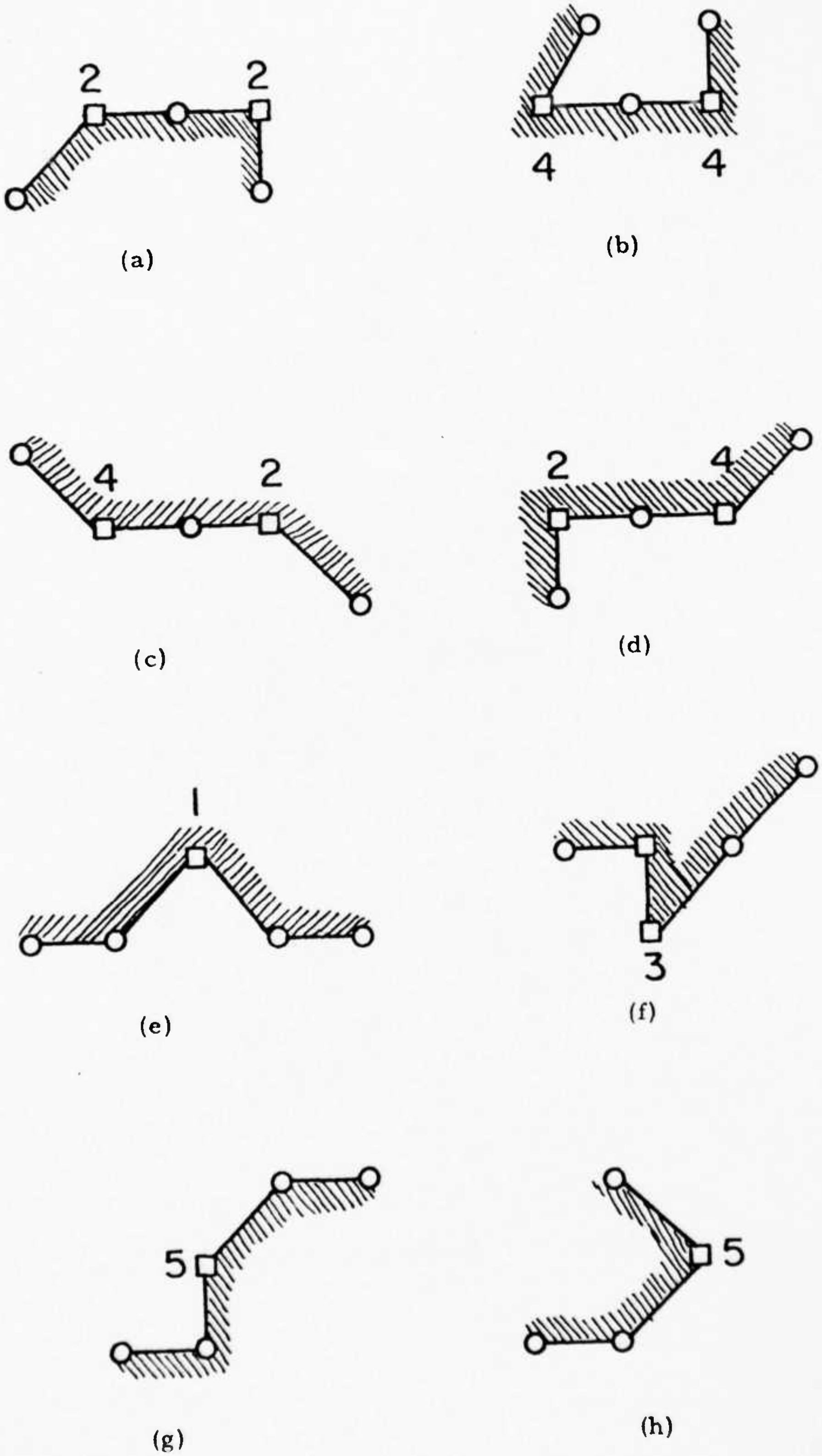


Figure 4. Typical combinations of types.

(a) types 1 or 3 (figure 4f) and matched pairs of types 2
(figure 4a) are convex boundary segments which are ignored;
(b) type 5 (figure 4g) and unmatched pairs of 2 and 4 (figure 4c
the beginning of an interior row.

(2) If it is approached from the interior, then
(a) types 1 or 3 (figure 4e) and matched pairs of types 2
(figure 4b) are concave boundary segments which signif
continuation of an interior row;
(b) type 5 (figure 4h) and unmatched pairs of 2 and 4 (figure 4d
the end of an interior row.

These cases are easily distinguished by using two parameters, on the interior/exterior condition and one for remembering w horizontal boundary segment is scanned. Figure 5 shows the resu applying step 2 to the boundary types in figure 3.

The preceding algorithm requires only the list of bo coordinates. This could be the output of a region analyzer. computationally quite efficient since each boundary point is ex only once during each of the two steps. Storage requiremen minimal and are determined by the region size. The resulting p matrix can be mapped into any desired area for the purpose of synthesis.

The silhouette generator by Dudani <1> is similar ir respects. It operates on a list of boundary points together list of predecessor/successor types (for a discussion of these

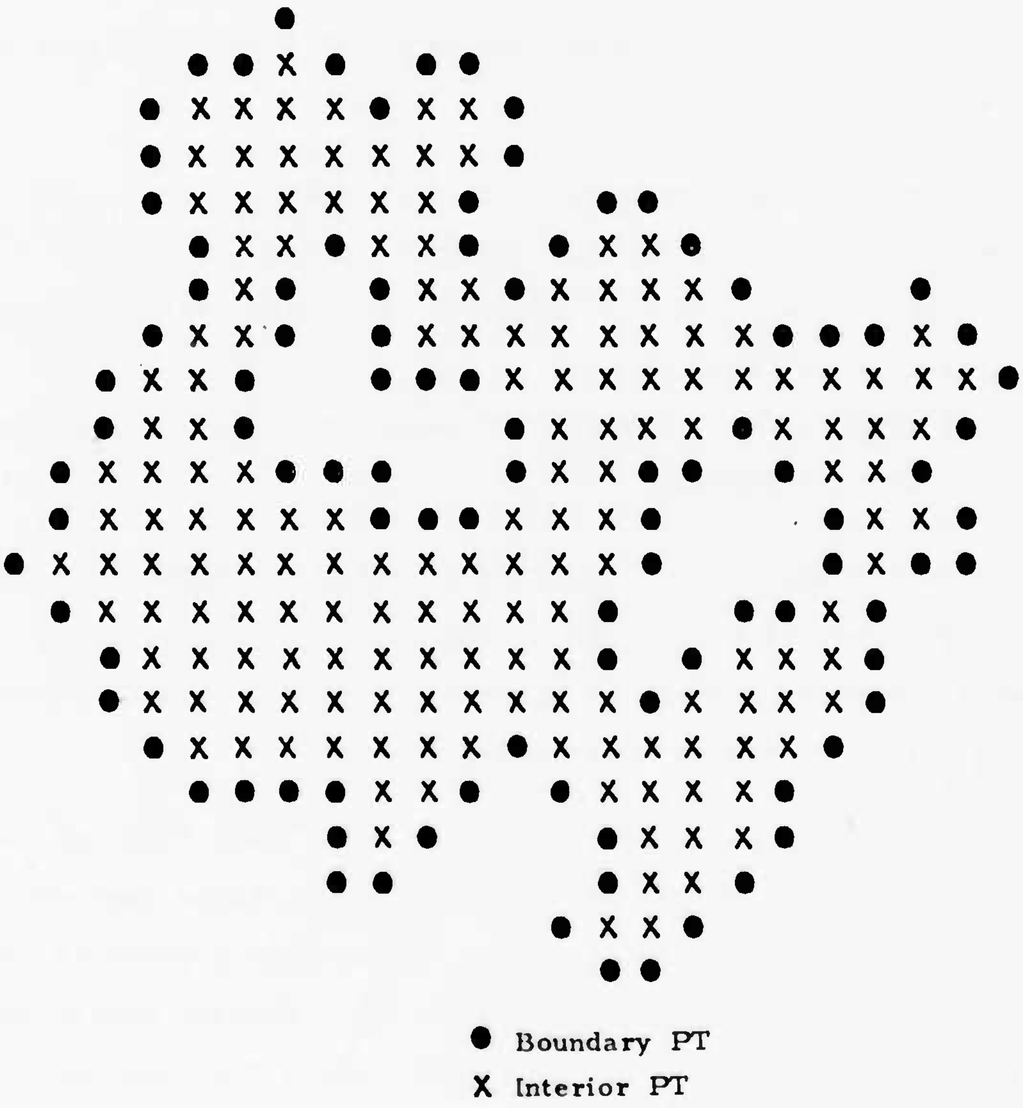


Figure 5. Generated region.

section 2.6 of this report). The notion of a "region crossing poin includes the above boundary type 5 as well as some other cases. T algorithm requires one additional step and cannot be used directly regions with holes.

Reference

1. S.A. Dudani, "Region Extraction Using Boundary Following," appear in Proceedings 1976 IEEE Workshop on Pattern Recognition a Artificial Intelligence.

2.9 AERIAL PICTURE SYNTHESIS

Benham Ashjari

Image synthesis involves the derivation by computer of "bas information" from an image and applying it to recreate the pictu from a set of available pictorial data bases.

The research currently being conducted involves digitized aeri photographs. In some cases, the synthesized version of a picture a the picture itself are close enough to describe various applicatio of image synthesis. In general, the synthesized picture h statistical and visual similarities to the original one. The "bas information" derived from the original image suffices to describe a form the synthesized version, and therefore this "basic information the information which represents the picture of interest and can stored or transmitted instead of the original picture. The techniq involves segmenting an area into distinct sections by texture or features, and preserving only the outline of the sections. The displaced segments are then compared with a predetermined set of data bases, and those most similar to the original sections are chosen for reproduction.

It can be shown that for transmission of the outline of a picture and synthesizing it at the receiving station we need 0.1 to 0.2% (1000:1 or 500:1) of the information bits that would be required to transmit a picture in the usual way.

Aerial Photographic Data Base

The data base used here is a set of 128 x 128 aerial pictures taken from a flight from Los Angeles to New York (see figure 3). There are nine distinct scenes in the data base:

1) Airport
2) Cloud patterns
3) Desert
4) Farmland
5) Forest
6) Mountain
7) Plains
8) High density urban areas
9) Water

These scenes are chosen to be mutually exclusive (i.e. such that they do not have other types of texture or scenes in them). For

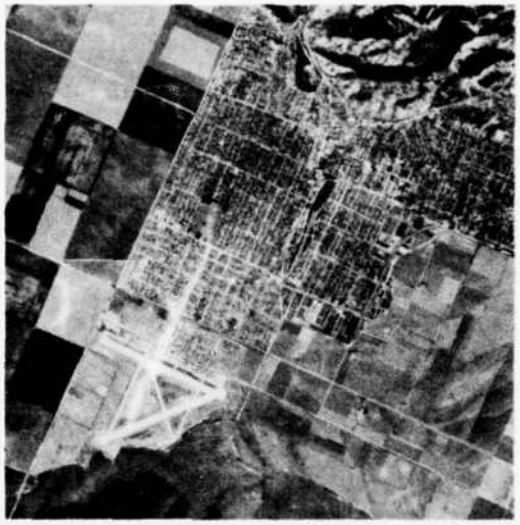


Figure 1. Original Picture

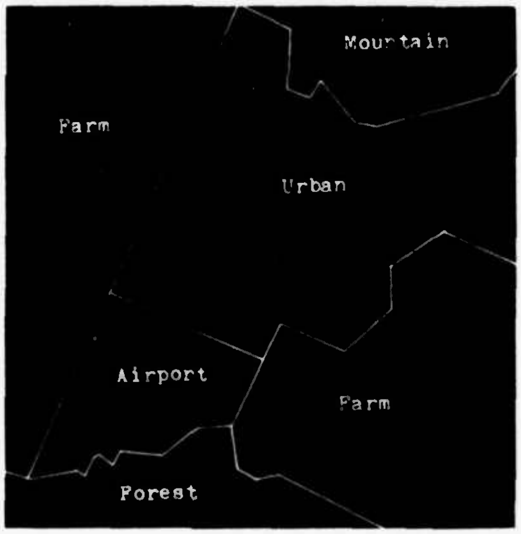
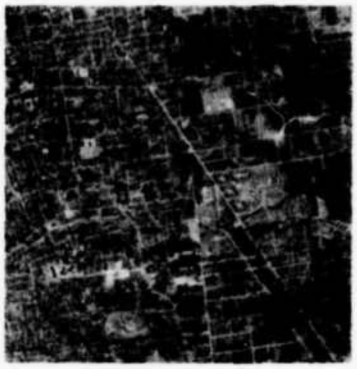
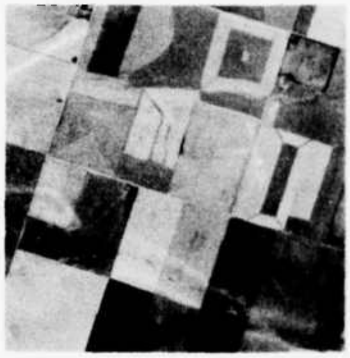


Figure 2. Segmented picture

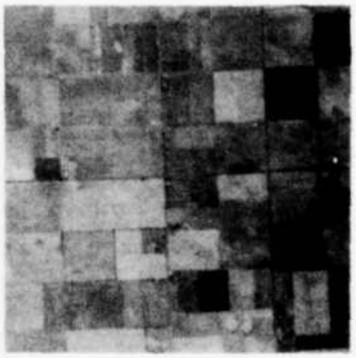
Urban file (09)



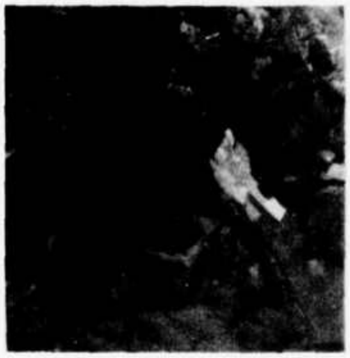
Urban file (13)



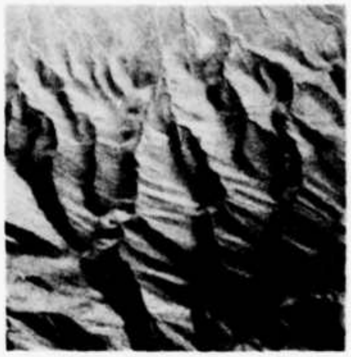
Farm file (37)



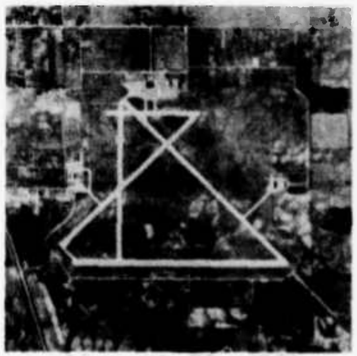
Farm file (48)



Forest file (54)



Mountain file (65)



Airport file (03)

Figure 3. Aerial data base used for synthesis of figure 1.

example, the urban area does not include a river passing through i Figure 3 shows some of the 128 x 128 data base pictures. Ideally need 32 pictures of each aerial scene (for a 5 bit code), and tl would make the total number in the aerial data base 288. However, have a total of 91 now available.

Statistical Analysis on the Data Base

In order to have a measure of closeness between the data ba pictures and the corresponding segmented area of the original pictur we need to set some mathematical criteria. A variety of measures ha been considered for coarseness, dissimilarity, homogeneity, a orientation which all fit into the area of texture analysis.

Gray-Level Co-occurrence Matrix

Before analyzing different criteria, we need to discuss t concept of gray-level co-occurrence matrix. (In the literature th is sometimes referred to as the grey tone spatial dependence matr <3>.)

Consider the (x,y) plane and a picture $f(x,y)$ defined over th plane. Let $\Delta = (\Delta x,\Delta y)$ be a vector in the plane. We can ccmpute t joint probability density of any pair of gray-levels separated by t vector Δ. For the discrete case, $\Delta x,\Delta y$ are integers. By counting t number of times that each pair of gray levels occur at separation Δ we can form an array of m x m, where m is the number of gray leve present in the picture. In this case, we count pairs of gray leve at separation Δ or $-\Delta$ indifferently. Therefore a symmetric matrix generated. For example, consider a 5 x 5 picture as the following:

$$\begin{array}{ccccc} 1 & 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 & 0 \\ 0 & 1 & 0 & 0 & 3 \\ 3 & 2 & 1 & 2 & 2 \\ 2 & 3 & 2 & 1 & 0 \end{array}$$

There are four gray levels 0,1,2, and 3 in the above picture. Let $(\Delta x, \Delta y) = (1,0)$. This means the pairs of gray levels in the x (or -x) direction and with a distance 1 from each other are counted. For all possibilities the gray level co-occurence matrix will be:

$$C_{\Delta} = \begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 2 & 3 & 1 & 1 \\ 1 & 3 & 2 & 5 & 0 \\ 2 & 1 & 5 & 4 & 6 \\ 3 & 1 & 0 & 6 & 0 \end{array}$$

We generally consider the vectors to be: (1,0), (1,1), (0,1), (-1,1). This corresponds to the pairs of pixels with 0°, 45°, 90° and 135° from the x-axis. Therefore, there are four (m x m) matrices for each (n x n) sample window taken out of the data base (m is the number of gray levels, n is the number of pixels in one row of the window). At the moment n is chosen to be 32 and we scale the gray levels of the pictures to be 32 levels instead of the normal 256.

Texture

Texture analysis plays an important role in image synthesis. To be able to replace a part of the original picture which includes, for example, farmland, with one of the 32 types of farmland pictures in the data base, we have to compare the textures in these two scenes and pick the closest one.

If a texture is coarse, and Δ is small ccmpared to the sizes the texture elements, the pairs of points at separation Δ shou usually have similar gray levels. This means that the high values the matrix C_Δ should be concentrated on or near its main diagona Conversely, for a fine texture, if Δ is comparable to the textu element size, then the gray levels of points separated by Δ shou often be quite different, so that the values in C_Δ should be spre out relatively uniformly. Thus a good way to analyze textu coarseness would be to compute, for various values of the magnitude Δ, some measure of the scatter of the C_Δ values around the ma diagonal. In what follows P(i,J) is the (i,J) th element of the giv matrix (which has size m x m = 32) divided by a normalizing factor which is equal to the sum of all of the matrix entries <4>. P(i,J) actually a measure of joint probability density that the pairs of gr levels (i,J)'s occur at pairs of points separated by $\Delta = (\Delta x, \Delta y)$.

Measure of correlation:

$$f_1 = \sum_{i=1}^{m} \sum_{J=1}^{m} i \cdot J(\frac{P(i \cdot J)}{R})$$

f_1 is a measure of coarseness <2>. f_1 can be normalized using mea and standard deviations associated with the marginal density functio $P_x^{(i,J)}/R$, $P_y^{(i,J)}/R$.

Homogeneity measures:

$$f_2 = \sum_{i=1}^{m} \sum_{J=1}^{m} (\frac{P(i,J)}{R})^2$$

(sum of squares of all matrix entries) f_2 is the angular second mome and is smallest when the P(i,J) are all as equal as possible. It large when some values are high and others are low <3,4>.

$$f_3 = \sum_{i=1}^{m} \sum_{J=1}^{m} P(i,J)/R \log(P(i,J)/R)$$

f_3 is the entropy and is another measure of homogeneity in a picture. This measure is largest for equal P(i,J) and small when they are very unequal.

Other than the above, the radial distributicn of Fourier Power Spectrum gives a measure of coarseness but was ruled out because it is believed to have a high computational cost and poor performance compared to the other textural features.

Orientation

Orientation is actually another measure of coarseness and it shows that a texture is coarser in one direction than another. Orientation is important in image synthesis because of the directionality of some natural patterns, e.g., mountains or man-made scenes such as urban areas or farmlands.

If a texture is directional then the degree of spread of the values about the main diagonal in C_Δ should vary with the direction of Δ (assuming that the Δ magnitude is in the proper range). This texture directionality can be analyzed by comparing spread measure of C_Δ for various directions of Δ <4>.

A measure for directionality has been considered to be the contrast:

$$f_4 = \sum_{i=1}^{m} \sum_{J=1}^{m} (i-J)^2 \, P(i,J)/R$$

This is essentially the moment of inertia of the matrix around its main diagonal; it is a measure of the degree of spread of the matrix values. It is expected to have the highest value at the direction perpendicular to the directionality of the texture.

Also the "angular" distribution of values of Fourier Power Spectrum is sensitive to the directionality of the texture in f. It will have its high values concentrated around the perpendicular direction to the directionality of the texture. However, at this moment it is not being considered for our measurements.

Feature Vectors

A 32 x 32 window is taken from each data base picture and a gray level co-occurrence matrix is computed for each of the four directions 0°, 45°, 90° and 135°. f_1, f_2, f_3 and f_4 are computed for each matrix. There are, therefore, 16 numbers associated with each data base picture. These 16 numbers can be assumed to be the coordinates of a 16-dimensional vector. Therefore, there are 32 vectors for each class (scene) of data base $(D_{i1}, D_{i2}, \ldots D_{i32})$ and we have nine classes in the data base $(1 \leq i \leq 9)$. These numbers will be saved on tape for future use.

Segmentation

The first stage in synthesizing a picture is delineating th distinct areas and sectioning them according to their texture an features (see figures 1 and 2). (It should be mentioned that because of unavailability of a simple picture which could include various textures and scenes, the original data (figure 1) was composed such that it included various desirable pictorial data.)

Segmentation is achieved by using the track-ball on the digital interactive display (Aerojet)* driven by the PDP 11/40. The locations of the pixels over which the track-ball passes is determined and stored in a 2000 x 2 array. The first column of the array corresponds to the row-locations and the second column of the array gives the column location of the pixels. The row and column location of each pixel is put in the array as the track-ball passes over them. They are the pixels which are located on the contours of segmented pictures. This 2000 x 2 array along with the description of what are inside the contours (mountain, forest, urban, etc.) are sent to the user's directory in a PDP KI10 computer on the ARPANET.

The user can now start synthesis of the picture. The synthesis is done automatically. The user should answer some questions asked by the program on his terminal and the operator mounts the data base tape and the feature vectors tape.

Statistical Analysis on the Original Picture and Selection of Appropriate Data Base

*An automatic segmentation routine based on pattern recognition techniques is being developed such that any picture may be segmented by computer rather than manually.

Once the different areas have been isolated (either manually c automatically), each must be matched against a wide selection c standard data bases. For example, there are 32 types of mountains i the data base. The mountainous area in the original picture must b replaced with the closest of these 32. For this purpose, we have t take a window out of the scene to be synthesized and compute the gra level co-occurrence matrix in four directions and basically do th same type of calculations as described earlier and come up with 1 numbers which constitute a feature vector Z. Suppose the featur vector is from the mountain class. We have to compare each vecto $D_{mountain1}$, $D_{mountain2}$, ..., $D_{mountain32}$ with Z and pick the closest No scheme has been developed yet for this purpose. However, th Fisher linear discriminant technique has been discussed in th literature and may be considered. We can use a voting scheme t determine the closest data base for a given measurement Z.

Synthesis

The receiving computer has the following information: th segmentation grid and the data base numbers with their respectiv positions. Now the appropriate data base is selected for eac section. If the computer finds the data base smaller than the spac it is to fill, it automatically magnifies it 2 x 2, using a interpolative scheme (see figure 5).

The computer then matches the grid section to the data base cutting off marginal areas, and pieces all of the segments togethe What results is a composite picture with great similarities to th original (see figures 4 and 5).

Because the project is at such an early stage, the automatic selection of data base by computer was not possible. The data bases in figures 4 and 5 are chosen by the user for illustrative purposes. The choices are according to visual similarities to the areas in the original picture.

The problem faced is that in a synthesized picture roads and rivers may appear discontinuous and only by extensive research in the future can we hope to overcome this problem.

Application

For any n x n picture the total number of points goes as n^2, while the number of points on the edges is proportional to n.

For transmission of a 256 x 256 picture which will have values from 0 to 255, there is a need for 256 x 256 x 8 = 524288 bits of information.

If this picture can be synthesized, the amount of information bits needed for transmission will be substantially reduced. This involves segmenting the picture to its differentiable areas, such as mountains, farmland, etc. This task is achieved by drawing a contour around each area. The result is a grid. For transmission of the location information of the contour points, on the grid, a great deal of advantage arises due to their connectivity. The beginning points of a contour requires 16 bits to specify its location. Then, through

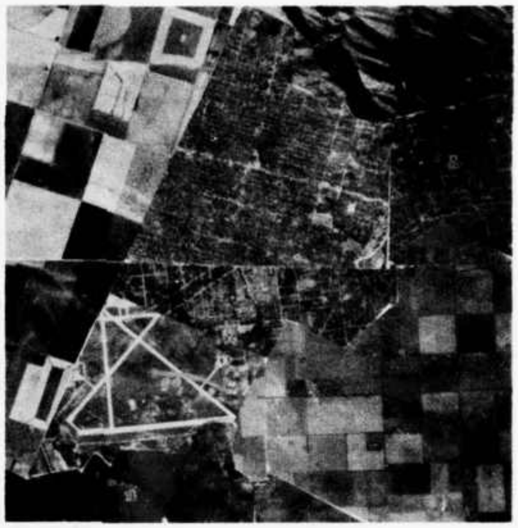


Figure 4. Synthesized Picture of Figure 1

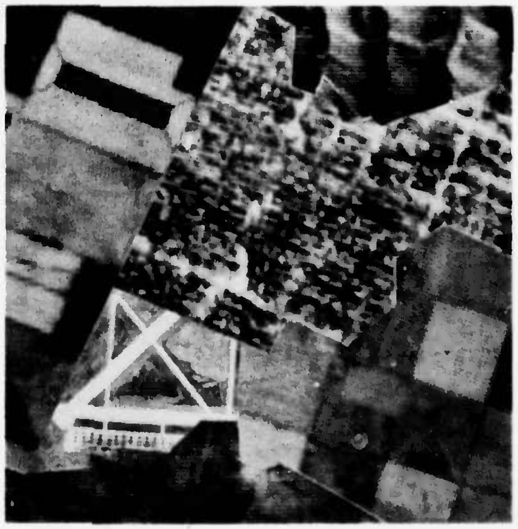


Figure 5. Synthesized with 2 x 2 magnification

The above figures (4 and 5) represent the simplest form of Image Synthesis for figure 1.

a chain coding process we can specify the seccnd pcint of the contc according to its local difference with the first one and so on. can approximate small variations on the contour with straight lii and therefore we have to code only the important turning points on t contour and with this procedure we probably need less than 60 poin to be coded. Considering the average distance of two points to be pixels, we need 8 bits for each of them, and we have 9 differe scenes with at most 32 variations on each one in the data bases. Th takes about 9 bits of information for each area and supposing the are 10 segmented areas on a picture; for transmitting th information, we need 586 bits which is between 0.1 to 0.2% of t information needed to transmit the entire original picture. This 99 to 99.9% saving has some applications. For example, in some case the local relations and types of scenes in a picture are of concer not the actual picture. For this case, image synthesis shall useful.

In some occasions that the actual data is needed in part of picture, synthesis is detrimental for that part. For these occasi that part must be transmitted exactly and the rest can be synthesiz For example, in figure 1, the airport can be transmitted exactly the rest of the picture can be synthesized.

In meteorological applications, original cloud patterns can retained while the exposed surface features are synthesized. synthesized picture will be an exact replica and there will st remain many cases in which the actual photographs are necessary.

these, image synthesis will never be useful. But for a great many uses, the time and cost saved in transmission of image data compensates for the loss of originality of the picture. The diagram shows various stages of synthesis and transmission.

References

1. H.C. Andrews, "Clustering for Image Segmentation," USCIPI Report 660, March 1976.

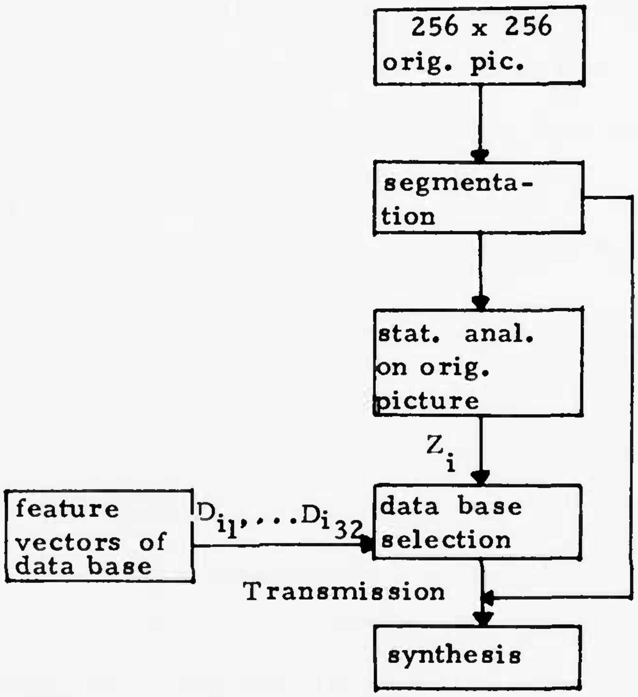
2. R.P. Kruger, "Image Textural Discrimination," USCIPI Report 444, February 1973.

3. R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, November 1973.

4. J.S. Weszka, C.R. Dyer, A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-6, No. 4, April 1976.

5. E.B Troy, E.S. Deutsch, and A. Rosenfeld, "Gray-Level Manipulation Experiments for Texture Analysis," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, January 1973.

6. D.N. Graham, "Image Transmission by Two-Dimensional Contour Coding," *IEEE Proceedings*, Vol. 55, No. 3, March 1967.


256 x 256
orig. pic.
segmentation
stat. anal. on orig. picture
Z_i
feature vectors of data base
$D_{i_1}, \ldots D_{i_{32}}$
data base selection
Transmission
synthesis

The above diagram shows various stages of synthesis and transmission.

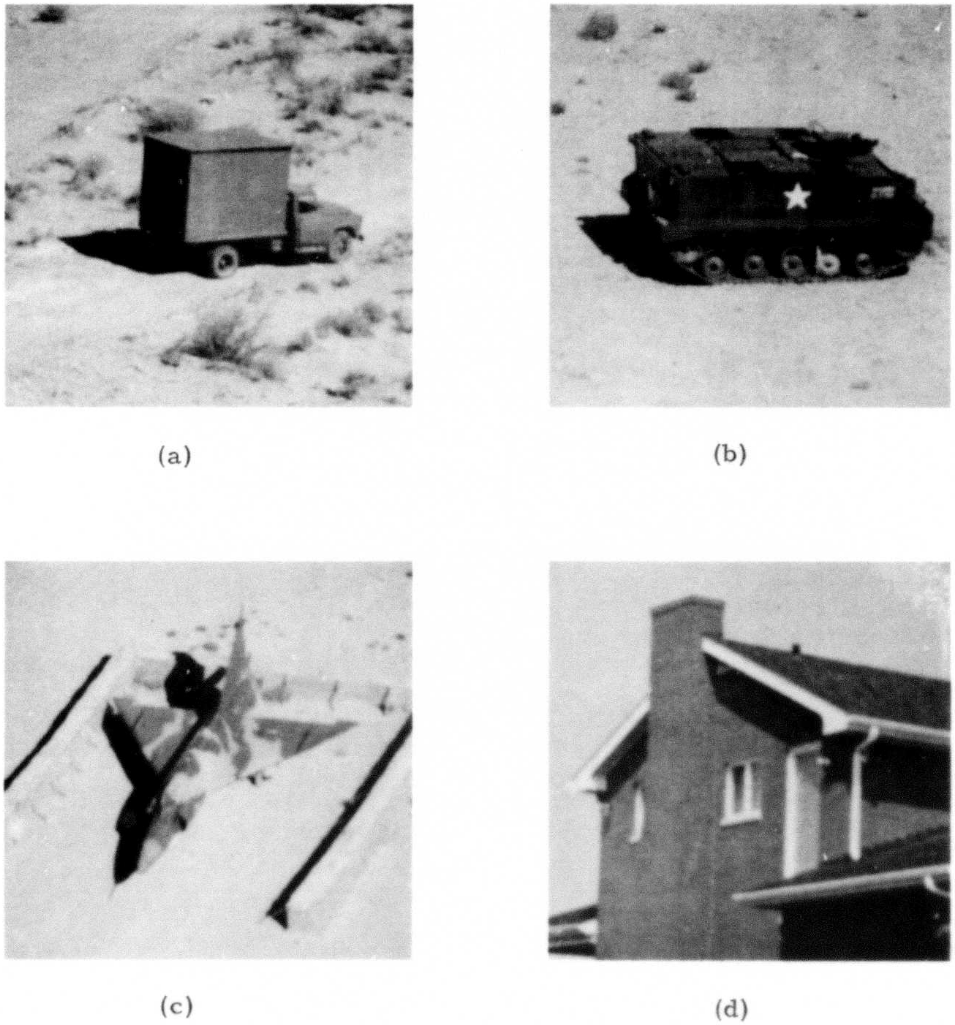
2.10 SOME EDGE SEGMENTATION RESULTS

Ramakant Nevatia

Satisfactory segmentation of objects in a scene is a key part of most scene analysis systems. The strategies employed for segmentation vary along many dimensions, often determined by the particular domain of the problem. Here, we restrict to the class of problems where no prototype specific information about the scenes is available, i.e., the objects present are not restricted to a small set of objects, nor any information known about their spatial relationships.

Results for four scenes, shown in figure 1 are presented here. All four pictures contain only a few (incidentally only one each) objects of interest and the objects are large (occupy a significant proportion of the picture). The first three pictures are of military vehicles against a desert background. The chosen pictures represent perhaps the simplest class of pictures, useful for practical applications. Yet, by the current state of the art, segmentaticn of these pictures is a difficult task (at least, without using prior knowledge). The complicating factors are the presence of texture and lack of multi-sensory information, such as cclcr or range.

Partial segmentation results using an edge based approach are presented here. Figures 2a through 2d show the edges detected in the corresponding pictures by the application of a Hueckel Edge Cperator <1>. The operator was applied at every seccnd pixel to every other row. A small vector along the direction of the edge is displayed for each detected edge. Note that same thresholds were used for all four



(a) (b)

(c) (d)

Figure 1. Four selected scenes.

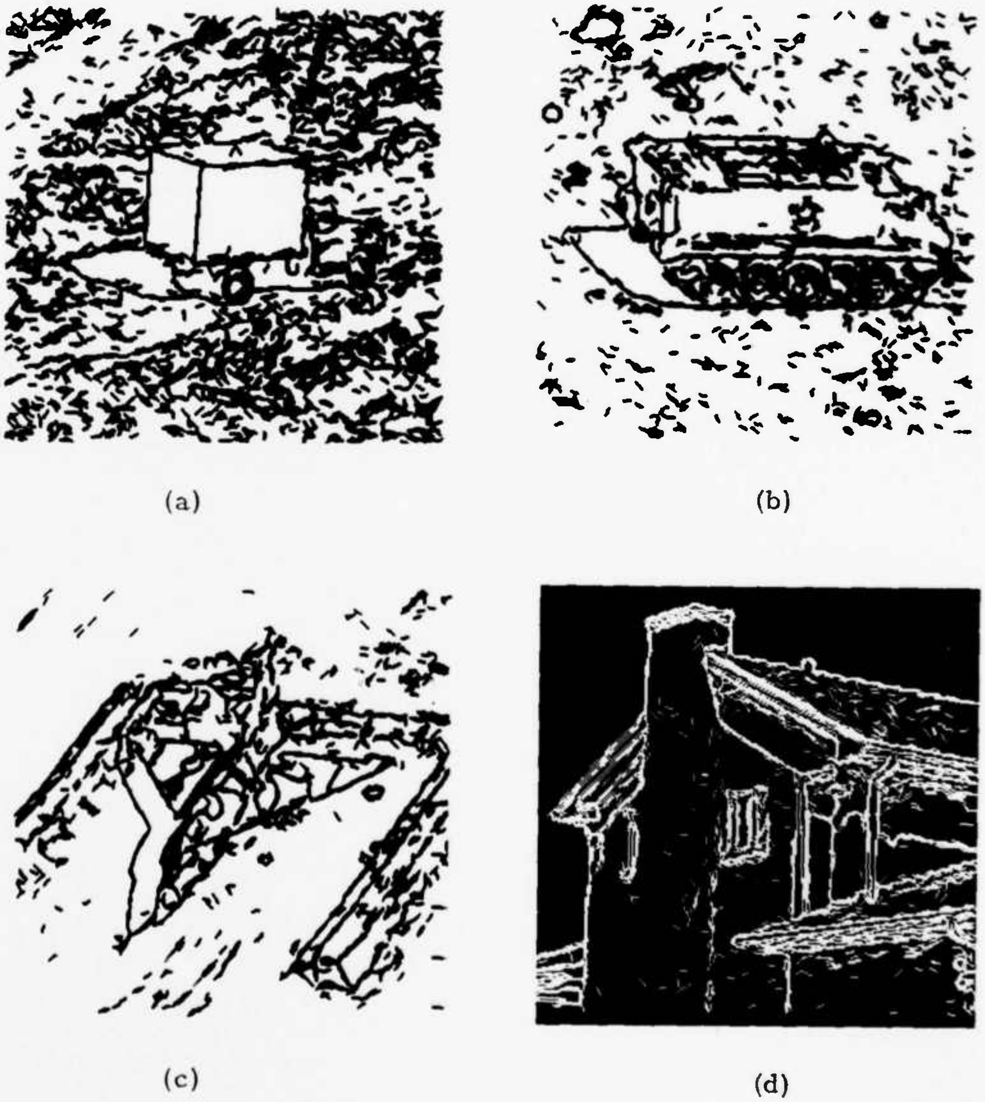


Figure 2. Edges detected in pictures of Figure 1.

pictures and no manual tuning for optimum edge detection wa attempted.

The edge pictures contain a large number cf edges belonging t the textured background but the objects are clearly visible to th human eye. It is hypothesized that edges belonging to object (particularly man-made objects) tend to occur along elongate segments, whereas most background texture edges tend to be distribute randomly. An edge linking procedure that links edges with simila orientations in straight line segments along chosen directions ha been described previously <2>. Results of applying this linkin procedure to edges shown in figure 2 are shown in figure 3. Here only those edge segments which contain at least seven edge element are retained.

Of course, such linked edge segments do not constitute a complet segmenation of the scene. However, they offer promising startin points for further segmentation. The next step, for example, may be completion of the partial boundaries of the cbjects. For man applications, such as recognition, the partial boundaries may suffic in many cases. Research is in progress for further extending thi approach.

Results shown here indicate that the techniques used are fairl successful in isolating extended edges of objects of interest. Mar of the difficulties arise in the process of edge detection itsel Hueckel edge detector is unable to detect fine, high resolution edg because of the large size of the operator and also fails in presen

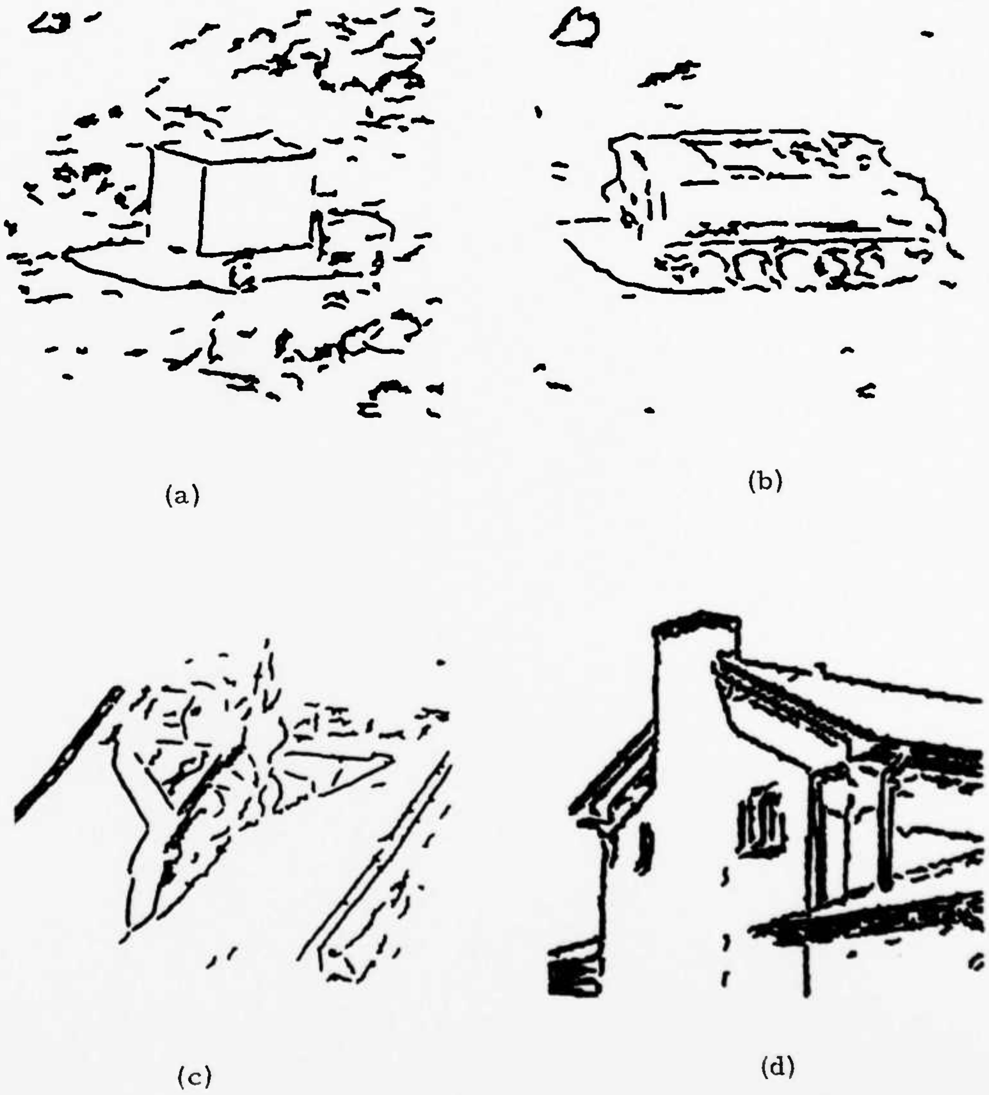


Figure 3. Linked edge segments from edges of Figure 2.

of fine texture. Local high resolution edge detectors, and al larger, texture edge operators may be used to supplement the edg detected by Hueckel operator.

A major motivation of presenting these intermediate results is stimulate comparison of techniques used by various groups working the ARPA Image Understanding program. It is felt that the select scenes are good test vehicles for segmentation techniques. (A larg number of similar pictures is now available in the data base prepar by USC.)

The picture used in figure 1d is a sub-picture of the house sce used in Ohlander's work at CMU <3>. Ohlander's technique relies recursive segmentation based on histogramming of various ima attributes. For the house scene, the most effective attributes se to be those based on color, such as hue and saturation. Note that t results presented here make no use of color information at all. (U of color in segmentation using techniques similar to those describ here is discussed in <4>.)

References

1. M. Hueckel, "An Operator which Locates Edges in Digiti Pictures," Journal of the ACM, Vol. 18, No. 1, January 1971, p 113-125.

2. R. Nevatia, "Object Boundary Determination in a Textu Environment," Proceedings of the ACM '75 Conference, Minneapol Minnesota, October 1975, pp. 32-36.

3. R. Ohlander, "Analysis of Natural Scenes," Ph.D. Thesis, Computer Science Department, Carnegie-Mellon University, August 1975.

4. R. Nevatia, "A Color Edge Detector," to appear in the Proceedings of the Third International Joint Conference on Pattern Recognition, November 8-11, 1976, Coronado, California.

2.11 EDGE DETECTION THROUGH ORTHOGONAL TRANSFORMATICN

Ikram E. Abdou

The problem of edge detection was discussed through many different approaches <1>. However, so far, the application of orthogonal transformation in edge detection has not been given enough consideration. One of the few examples in this trend is the Hueckel operator <2>. In Hueckel's work edge detection is based on fitting the intensities in a given region to an ideal edge element. The position and direction of this edge is obtained by minimizing the mean square error between the intensities of the region considered and those of the ideal edge. Computation is simplified by transforming into the Fourier domain and using only the first eight coefficients. A measure of the goodness of edge is computed and is used in accepting or refusing the edge. His work was extended to line detection in <3>.

Although Hueckel's results are quite good, the operator has two disadvantages. First, it is time consuming and second, small edges cannot be detected in the relatively large block size used.

Another approach to the problem will be introduced here. In t approach a finite set of edges in a small block is considered. corresponding Fourier coefficients are computed. Edge detection implemented by comparing these coefficients to suitable thresholds.

The computation of Fourier coefficients is a simple problem, can be easily generalized to shifted or rotated edges. This is the case for the detection problem, where a statistical model needed to obtain an optimum detection strategy. In fact the w described in <4> can be used in solving this problem.

The set of edges considered in this paper is shown in figure A function f(x,y) defined on an N x N block has corresponding Four coefficients F(u,v) given by

$$F(u,v) = \sum_{y=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{x=-\frac{N-1}{2}}^{\frac{N-1}{2}} f(x,y)\, W^{xu+yv} \quad * \qquad (1)$$

where

$$W = \exp\left(-\frac{2\pi Y}{N}\right) \qquad (2)$$

In many cases closed forms for the Fourier coefficients can derived. As an example, for the case of a horizontal edge given by

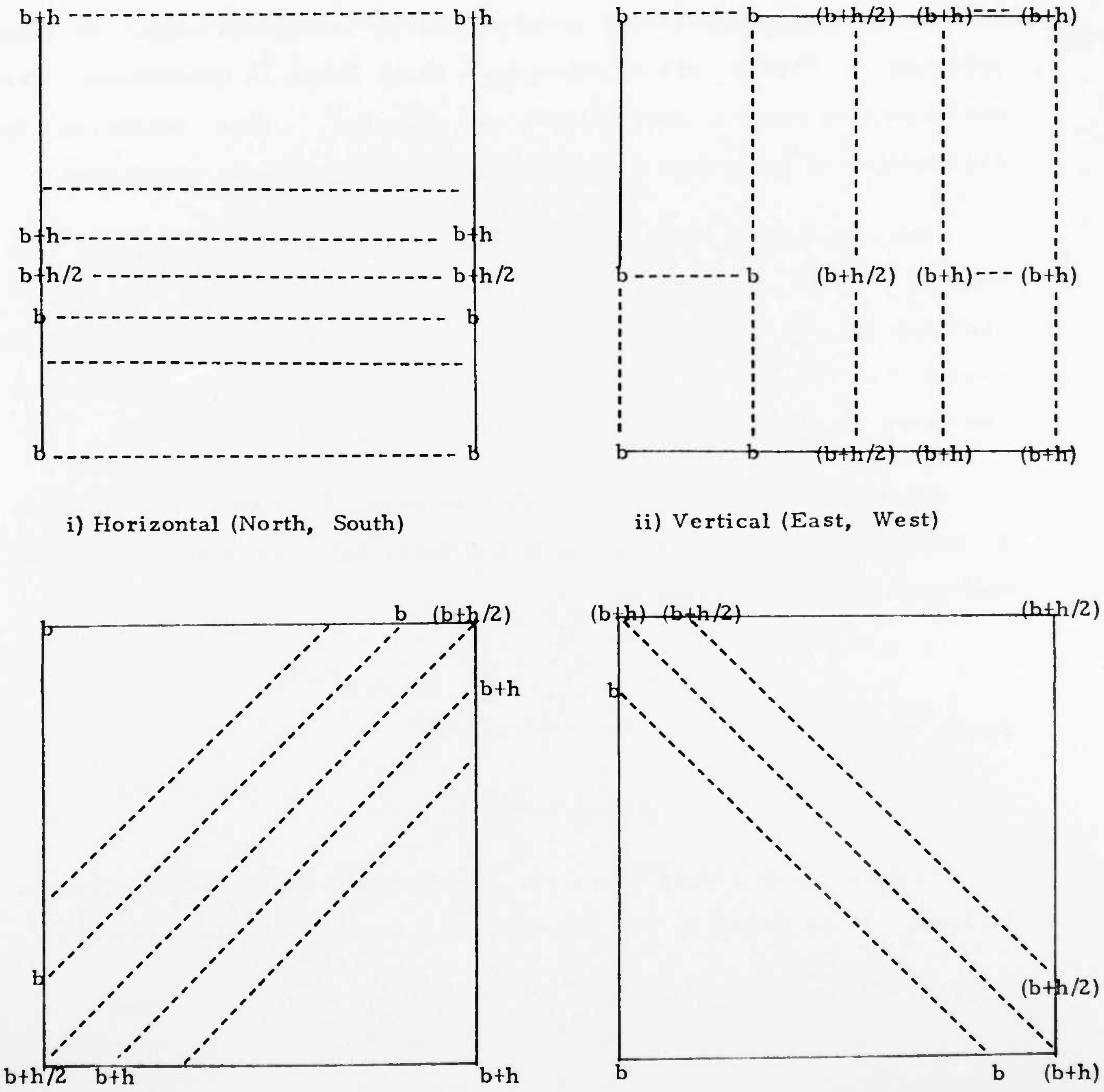
$$f_H(x,y) = b \qquad \text{for } y < 0 \qquad (3a)$$

$$= b + \frac{h}{2} \qquad \text{for } y = 0 \qquad (3b)$$

$$= b + h \qquad \text{for } y > 0 \qquad (3c)$$

The corresponding Fourier coefficients are

$*\ \frac{1}{N^2}$ is dropped for simplification.



i) Horizontal (North, South)

ii) Vertical (East, West)

iii) π/4 (Southeast, Northwest)

iv) -π/4 (Northeast, Southwest)

Figure 1. Basic edge elements f(x, y) for h (positive, negative)

$$F_H(0,0) = N^2(b + \frac{h}{2}) \qquad (4)$$

Otherwise

$$F_H(o,v) = Nh \quad \frac{1}{2} + \frac{W^{\frac{N+1}{2}v} - W^v}{W^v - 1} \qquad (5)$$

$$F_H(u,v) = 0 \qquad (\text{if } u \neq 0) \qquad (6)$$

Fourier coefficients are calculated for all the cases shown figure 1. The results for N = 5 are shown in figure 2.

From these results it is obvious that the edge direction can determined from the Fourier coefficients. This technique is used wi the test image given in <5>. The Fourier coefficients correspondi to each block are compared to different thresholds to decide t presence or absence of an edge. The direction of the edge, present, is also determined. Edge maps for SNR = 10 and 1 are sho in figures 3 and 4, respectively. In figures 3a and 4a only t information about edge location is used, while in figures 3b and information about the edge direction is used to enhance the edge ma In both cases the resulting edge maps show some degree of improveme over those obtained by the Sobel operator <5>.

References

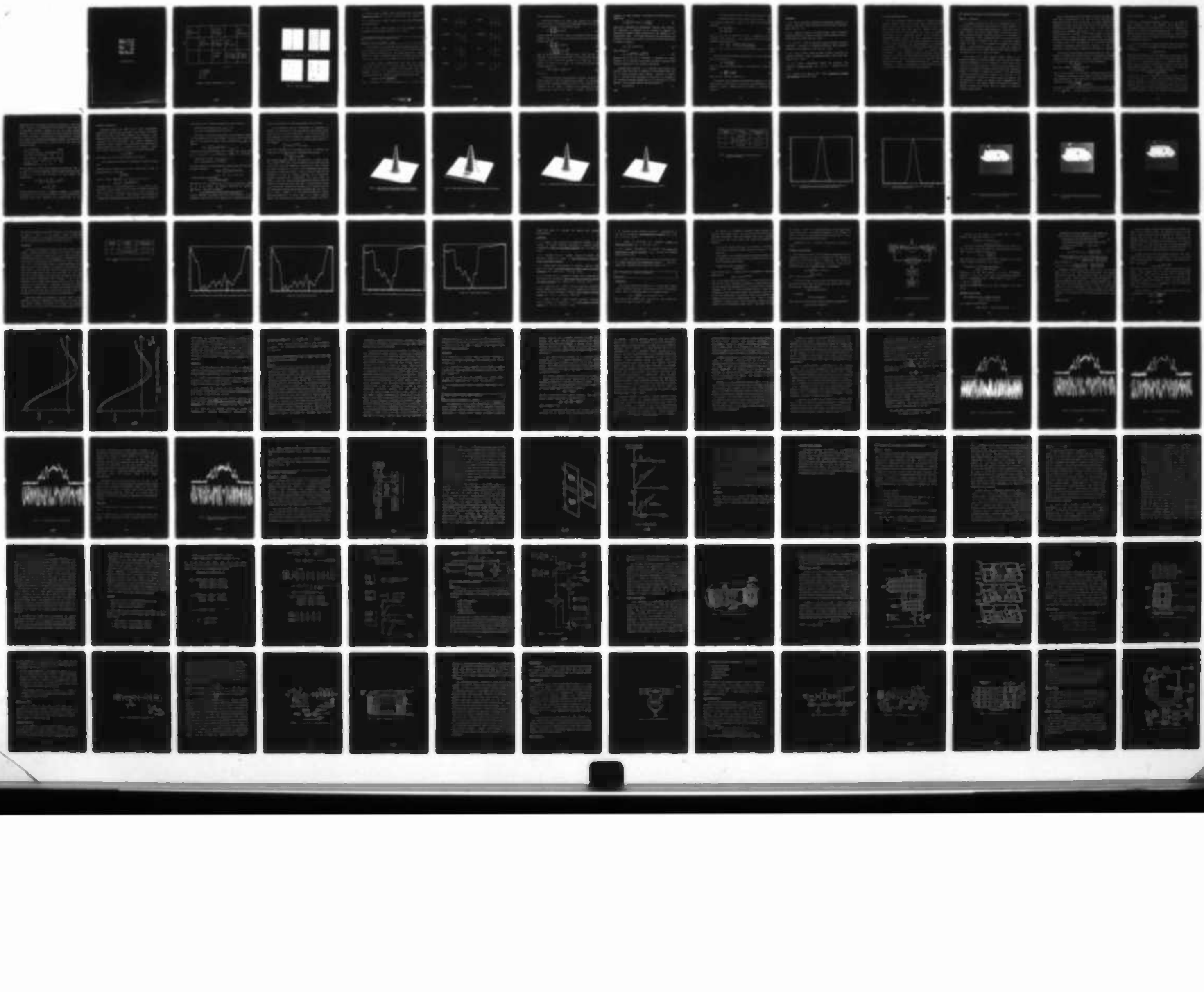
1. L.S. Davis, "A Survey of Edge Detection Techniques," Universi of Maryland Computer Science Center, November 1973.

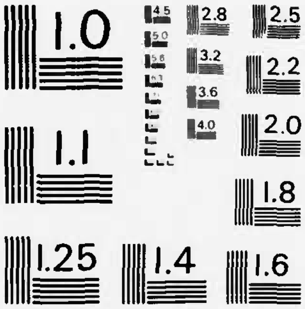
2. M.H. Hueckel, "An Operator which Locates Edges in Digiti Pictures," Journal of the ACM, Vol. 18, No. 1, January 1971,

AD-A134 943 IMAGE PROCESSING RESEARCH(U) UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES IMAGE PROCESSING INST H C ANDREWS 30 SEP 76 USCIPI-720 F33615-76-C-1203

2/3

UNCLASSIFIED F/G 20/6 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

$v \uparrow$

i) 0 ii) 0 iii) 0.812 jh iv) 0		i) 1.82 jh ii) 0 iii) -2.63 jh iv) 2.63 jh		i) 0 ii) 0 iii) 0 iv) -0.812 jh
	i) 0 ii) 0 iii) 3.44 jh iv) 0	i) -7.69 jh ii) 0 iii) 4.25 jh iv) -4.25 jh	i) 0 ii) 0 iii) 0 iv) -3.44 jh	
		(i, ii, iii, iv) $25(b+\frac{h}{2})$	i) 0 ii) -7.69 jh iii) -4.25 jh iv) -4.25 jh	i) 0 ii) 1.82 jh iii) 2.63 jh iv) 2.63 jh

$\longrightarrow u$

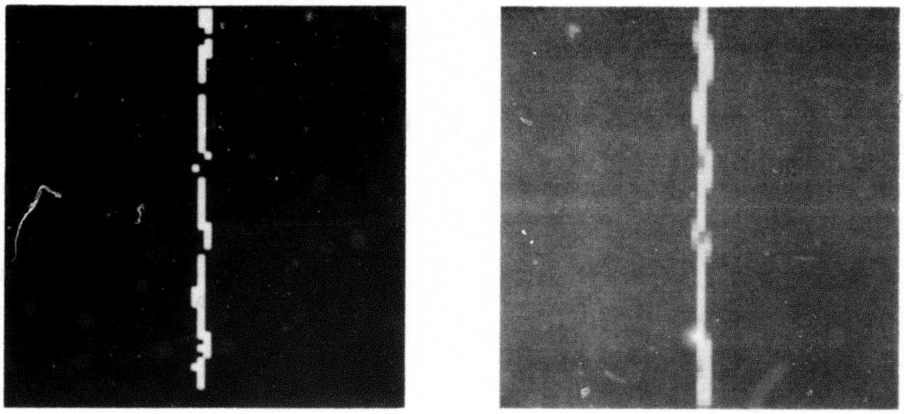
i) Horizontal

ii) Vertical

iii) $\pi/4$

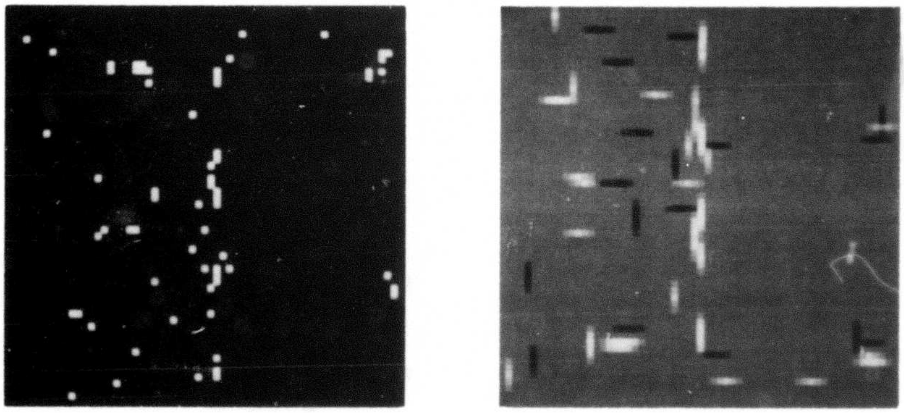
iv) $-\pi/4$

Figure 2. Fourier coefficients for a 5 x 5 block.



(a) (b)

Figure 3. Edge maps for SNR = 10.



(a) (b)

Figure 4. Edge maps for SNR = 1.

113-125.

3. "A Local Visual Operator which Recognizes Edges and Lines," Journal of the ACM, Vol. 20, No. 4, October 1973, pp. 634-647.

4. I. Abdou and W.K. Pratt, "Statistical Design of Edge Detectors," see present report.

5. W.K. Pratt, "Figure of Merit for Edge Location," USCIPI Semiannual Technical Report, March 1976, pp. 85-93.

2.12 STATISTICAL DESIGN OF EDGE DETECTORS

Ikram E. Abdou and William K. Pratt

A wide variety of edge detection methods have been devised <1-3>. Only recently has there been any attempt to devise a figure of merit for edge detector performance <4>. This report summarizes an ongoing effort to develop a statistical model for edge detection that can be used in the design and a priori evaluation of edge detectors <5>.

Edge Model: Figure 1 illustrates the basic model for a local ideal edge defined over a 3 x 3 block of pixels. These ideal edges are assumed to be subject to additive white Gaussian noise with standard deviation σ. The probability density at each 3 x 3 point in an observation array is therefore a Gaussian density of the form

$$G(a,\sigma) = [2\pi\sigma^2]^{-\frac{1}{2}} \exp\left\{\frac{-(x-a)^2}{2\sigma^2}\right\} \qquad (1)$$

where Y represents the observed pixel amplitude and a is its mean

North	h	h	h	South	0	0	0
	h/2	h/2	h/2		h/2	h/2	h/2
	0	0	0		h	h	h
Northeast	h/2	h	h	Southwest	h/2	0	0
	0	h/2	h		h	h/2	0
	0	0	h/2		h	h	h/2
East	0	h/2	h	West	h	h/2	0
	0	h/2	h		h	h/2	0
	0	h/2	h		h	h/2	0
Southeast	0	0	h/2	Northwest	h	h	h/2
	0	h/2	h		h	h/2	0
	h/2	h	h		h/2	0	0

Figure 1. Ideal Basic Edges

value as specified by figure 1.

Linear Edge Detection: In a linear edge detector the observed pixel array F(j,k) is multiplied by an enhancement mask M(j,k) over an I x L window to produce an edge gradient function

$$G = \sum_{j=1}^{L} \sum_{k=1}^{L} F(j,k)\,M(j,k) \qquad (2)$$

Since the gradient operation is linear, the probability density of the gradient $\rho(G) = \mathcal{N}_E(u,o)$ will also be Gaussian with mean and standard deviation given by

$$u = \sum_{j=1}^{3} \sum_{k=1}^{3} F(j,k)\,M(j,k) \qquad (3a)$$

$$\sigma = \sigma \left[\sum_{j=1}^{3} \sum_{k=1}^{3} M^2(j,k) \right]^{\frac{1}{2}} \qquad (3b)$$

where the overbar indicates an ensemble average. Edge detection is performed by thresholding the magnitude of the edge gradient function G such that $|G| \geq T$ implies the presence of an edge, and $|G| < T$ implies the absence of an edge. The probability density of the gradient magnitude is

$$p(|G|) = [\mathcal{N}_E(u,\sigma) + \mathcal{N}_E(-u,\sigma)] \qquad (4)$$

for $G \geq 0$.

The detection procedure outlined above is entirely analogous to signal detection in radar and communication systems. The performance of such systems is commonly measured in terms of the probability of detection when the signal is present, and the probability of false detection (false alarm) when no signal is present. Applying this technique to edge detection, the probability of false detection is found to be

$$P_F \int_T^{\infty} p(|G| \mid u=0)\,dx = 2\,\mathrm{erfc}\left[\frac{T}{\sigma_E}\right] \qquad (5)$$

where erfc[·] represents the complementary Gaussian error function. The probability of detection can be obtained by computing the probability of making a correct decision given the presence of the corresponding edge. If all edge orientations are equi-probable, the probability of detecting an edge of any orientation given that an edge is present is given by

$$P_D = \int_T^{\infty} p(|G| \mid u=u)\,dx \qquad (6a)$$

which reduces to

$$P_D = \mathrm{erfc}\left[\frac{T-u}{\sigma}\right] + \mathrm{erfc}\left[\frac{T+u}{\sigma}\right] \qquad (6b)$$

Equations (5) and (6) may be employed to determine the optimal threshold as a function of noise level to achieve a specified false detection probability or true detection probability.

Nonlinear Edge Detection: In a nonlinear edge detector the observed pixel array F(j,k) is combined in some nonlinear manner over an L x L window to produce an edge gradient. Analysis is possible only if the form of the nonlinearity is specified. As an example of the statistical procedure consideration will be directed to the Sobel operator. With the Sobel operator the edge gradient is

$$G = [X^2 + Y^2]^{\frac{1}{2}} \qquad (7)$$

where

$$X \equiv [F(1,3) + 2F(2,3) + F(3,3)] - [F(1,1) + 2F(2,1) + F(3,1)] \tag{8a}$$

$$Y \equiv [F(1,1) + 2F(1,2) + F(1,3)] - [F(3,1) + 2F(3,2) + F(3,3)] \tag{8b}$$

The probability densities of X and Y are given by the Gaussian densities

$$p(X) = \mathcal{N}(u_X, \sqrt{12}\,\sigma \tag{9a}$$

$$p(Y) = \mathcal{N}(u_Y, \sqrt{12}\,\sigma \tag{9b}$$

where u_X and u_Y are the means of X and Y, respectively, for a particular edge orientation. The probability density of the edge gradient is then found to be

$$p(G) = \frac{G}{12\sigma^2} \exp\left\{-\left(\frac{G^2 + u^2}{24\sigma^2}\right)\right\} I_o\left(\frac{Gu}{12\sigma^2}\right) \tag{10}$$

where $u^2 = u_X^2 + u_Y^2$ and $I_o(.)$ is the modified Bessel function of zero order.

With this statistical model the probability of false detection for no edge present is

$$P_F = \exp\left\{-\frac{T^2}{24\sigma^2}\right\} \tag{11}$$

where T is the gradient edge threshold. The probability of detection becomes

$$P_D = Q\left\{\sqrt{\frac{3}{2}}\ \frac{h}{\sigma}, \frac{T}{\sqrt{12}\,\sigma}\right\} \tag{12}$$

where Q(a,b) is Marcum's Q function <6>.

Conclusions: The statistical design procedure summarized in this report appears to hold promise for the design of edge detectors in the presence of noise. Further study is underway to evaluate the procedure.

References

1. W.K. Pratt, "Luminance Edge Detection Techniques," University of Southern California, Image Processing Institute, USCIPI Report 660, March 1976, pp. 4-21.

2. G.S. Robinson, "Detection and Coding of Edges Using Directional Masks," University of Southern California, Image Processing Institute, USCIPI Report 660, March 1976, pp. 40-57.

3. W. Frei and C. Chen, "A New Class of Edge and Feature Detection Operators," University of Southern California, Image Processing Institute, USCIPI Report 660, March 1976, pp. 22-40.

4. W.K. Pratt, "Figure of Merit for Edge Location," University of Southern California, Image Processing Institute, USCIPI Report 660, March 1976, pp. 85-93.

5. A.K. Griffith, "Mathematical Models for Automatic Line Detection," Journal of ACM, Vol. 20, No. 1, January 1973, pp. 62-80.

6. M. Schwartz, W.R. Bennett, and S. Stein, Communication Systems and Techniques, McGraw-Hill, 1966.

3. IMAGE PROCESSING PROJECTS

The image processing projects comprise an ongoing research activity directed toward image coding, image restoration, vision modelling, and the implementation of image processing systems. In image coding novel ideas based upon the results of the image understanding study are being explored as a means of achieving significantly higher compression ratios than obtainable by conventional coding methods. The image restoration studies are directed toward the solution of two major problems: blind restoration in which a priori information about image degradation is unavailable or incomplete; and constrained restoration which involves the use of luminance bounds and smoothness criteria to improved image restoration. Vision modelling research activities include the extension of previously developed models of the human visual system to encompass higher levels of visual perception in support of the image understanding program. Implementation studies are underway on techniques of nonlinear, two dimensional optical filtering which can be utilized as a form of sensor based image processing.

3.1 VARIABLE KNOT SPLINES AS AN IMAGE ANALYSIS TECHNIQUE

Dennis C. McCaughey

The degrees of freedom of an image or its information content is an important and frequently recurring subject in image processing. It is fundamental to the image coding problem and in other areas such as remote sensing. (Here remote sensing refers to underwater as well as RPV and satellite imaging.) In coding problems one is interested in the transmission of that information relevant to the users needs and in the elimination of irrelevant data, while in the remote sensing problem the image is to be obtained from samples of a corrupted version. In these areas the concern is the degrees of freedom of sampled images and in this context there have been several attempts at quantifying this notion of degrees of freedom.

With respect to the remote sensing problem, Twomey <1> has equated the degrees of freedom of samples output data with this number of non-zero eigenvalues of the gram-matrix of the linear system corrupting the original functions. This idea has been extended in two dimensions and applied to projection imaging systems with significant results by McCaughey and Andrews <2>. However, if the linear system corrupting the original function is unknown, as is the case in the blind deconvolution problem <3>, another approach is necessary. One such approach is to apply the singular value decomposition (SVD) algorithm <4> to the sampled image matrix whereupon the number of degrees of freedom can be equated with the number of effectively non-zero singular values.

Both of these methods deal with sampled versions of the image on some linear transformation of it, and are thus affected by the sampling method used. In the SVD case this is readily apparent by considering an image $f(x,y)$ that can be written as the product of two functions $f_1(x)$ and $f_2(y)$ as $f(xy) = f_1(x)\ f_2(y)$. If this image is sampled on a Cartesian grid (x_i, y_k), $i = 1,2,\ldots,N$, $k = 1,2,\ldots,N$. The image matrix can be written as the outer product of the two vectors $[f_1(x_1)\ldots f_1(x_N)]$ and $[f_2(y_1)\ldots f_2(y_N)]$ which is a rank one matrix for all such separable images f. The point is that the degrees of freedom should be a characteristic of the original image and reflected in the sampled image only by our inability to collect an uncountably infinite number of samples for application on a computer.

This brings up some conceptual difficulties since the number of degrees of freedom of a function defined on a continuum is countably infinite at best viz the space of all square integrable functions on $[-1,1]$, $L^2(-1,1)$, where any $f(x) \in L^2(-1,1)$ can be written as

$$f(x) = \sum_{i=1}^{\infty} f_i \phi_i(x)$$

$\{\phi\}$ being a complete orthonormal set and

$$f_i = \int_{-1}^{1} f(x)\phi^*(x)dx.$$

Here to exactly specify f in an L^2 sense a countably infinite set $\{f_i\}$ is required. In general if we are willing to accept an approximation with some error say, δ, then there exists a number $N(\delta,\varphi)$ which is a function of δ and $\{\varphi\}$, such that

$$\left[\int_{-1}^{1} \left|f(x) - \sum_{i=1}^{N(\delta,\varphi)} f_i \varphi_i(x)\right|^2 dx\right]^{\frac{1}{2}} < \delta.$$

We could then define the degrees of freedom of that f at level δ, DoF_δ as the $\inf\limits_{\{\varphi\}} N(\delta,\varphi)$. In this case where $f \in L^2(-1,1)$ for any complete set $\{\varphi\}$ we know that $\delta = \left[\sum_{i > N(\delta, \varphi)} |f_i|^2 \right]^{\frac{1}{2}}$

so that δ is a monotonically decreasing function of the number of basis functions used in the expansion. This property is of fundamental importance since without it, it would be impossible to compare two different orthonormal expansions for f. In what follows we will generalize this concept of epsilon degrees of freedom using one dimensional cases for simplicity and then extend the results to two dimensional expansions.

Let f be an element of normed linear vector space W with distance function d . Let $\mathscr{A}$ be an approximation scheme where the estimate $\hat{f}$ of f is given by

$$\hat{f} = \sum_{i=1}^{\eta} f_i \beta_i(x) \triangleq \mathscr{A}_n .$$

Here $N(\delta, \beta, d)$ is dependent upon δ , the sequence $\{\beta\}$ and the metric (d_N). We further require the property that $\forall \delta > 0$ there exists $N(\delta, \beta, d)$ such that $\eta \geq N(\ , \ , d)$ implies that $d_W(f, \mathscr{A}_\eta) < \delta$. Further we require that $N(\delta, \{\beta\}, d_W)$ be a monotonically increasing function of the precision to which we are approximating f. The epsilon degrees of freedom with distance d_W, $\mathrm{DoF}(d_W, \delta)$, can then be defined as follows:

$$\mathrm{DoF}(d_W, \delta) = \inf_{\{\beta\}} N(\delta, \beta, d_W).$$

To find the set $\{\beta\}$ is, in general, far too much to hope for. However in the case where W is taken to be the space of bandlimited L^2 functions observed over a finite subinterval of the real line, the distance d_W being the L^2 metric, the result is known. Here for every δ the $\inf_{\{\beta\}} N(\delta, \beta, d_W)$ was found by Landau and Pollack <5> to be achieved by the functions $\{\varphi\}$ which are related to the prolate spheroidal waveforms and first described by Slepian and Pollack <6>. While this result is known, the determination and utilization of these waveforms for large space bandwidth problems has met with little success. Therefore to arrive at any meaningful and useful results we will have to restrict the search to functions that are computationally feasible and possess the approximating properties required. One such set is the set of polynomial splines $S_{k,n}$ of degree k with n free knots and defined as follows

$S_{k,n} = \{s(t) \mid$ there exists
$0 = x_0 \leqq x_1 \leqq \ldots \leqq x_{r+1} = 1$ and integers $m_1 \ldots m_r$
with $1 \leqq m_i \leqq k+1$ for each i and $m = n$
such that s(t) Π in each of the intervals
(x_i, x_{i+1}) while $S \in C^{k-m_i}$ in an open
neighborhood of x_i, $i = 0, \ldots, r\}$.

Here Π_k denotes the class of all polynomials of degree less than k. $S_{k,n}$ thus is the class of polynomial splines allowing multiplicities up to order k+1 at each knot and every $s(t) \in S_{k,n}$ possess the representation

$$s(t) = \sum_{i=0}^{k} a_i t^i + \sum_{i=0}^{n} \sum_{j=0}^{m_i} c_{ij} (t - x_i)_t^{n-j+1}$$

where

$$(x)_+^n = \begin{cases} x^n & x \geqq 0 \\ 0 & 0^W \end{cases} .$$

Clearly the method of fitting a different pclyncmial in different subintervals of [0,1] as done by Pavlidis <7> can be classified as a pclynomial spline with appropriate knot multiplicities at the subinterval boundaries.

Polynomial splines are chosen due tc their approximation properties and the fact that they possess a basis namely, the B-spline basis, that provides a local basis property thus allowing a rapid generation while the matrices involved in generating a B-spline fit to a function f are well conditioned. With DeBoor's algorithm for computations using B-splines <8> no difficulties are encountered in handling multiple order knots. Hereafter we will consider a spline s(t) of order k with n knots $s(t) \epsilon S_{k,n}$ to be of the following form

$$s(t) = \sum_{i=1}^{n} b_i B_i^{(k)}(t)$$

where $B_i^{(k)}(t)$ is the i-th k order B-spline basis function.

In one dimension the following L^P approximation results are known <9>.

Theorem 1: Given f defined on [0,1] then there exists a best L_P approximation of f of the form

$$\sum_{i=1}^{n} b_i B_i^{(k)}(t).$$

Theorem 2 (DeBoor unpublished): Let f and n be fixed then if

$$\delta_k^2 = \int_0^1 \left| f(t) - \sum_{i=1}^{n} b_i B_i^{(k)}(t) \right|^2 dt \neq 0$$

where $\sum_{i=1}^{k} b_i B_i^{(k)}(t)$ is a best approximation of f in $S_{k,n}$ then $\delta_{k+1} < \delta_k$. Thus we can create a sequence $\{s_i(t)\}$ of best approximating splines of order k such that if $s_i(t) \epsilon S_{k,n}$, $s_{i+1}(t) \epsilon S_{k,n+1}$. Then $s_i(t) \rightarrow f(t)$ in the L^2 sense as i goes to ∞. Thus polyncmial splines possess the aforementioned approximating properties, and the degree of freedom at level epsilon in the L^2 metric, DoF$(k,\delta,\|\cdot\|_2)$, will be taken as

$$\text{DoF}(k,\delta,\|\cdot\|_2) = \min\{n \mid s_i(t) \in S_{k,n} \text{ where } s_i(t) \text{ is a best } L^2 \text{ approximation to } f \in L^2[0,1]\}.$$

The degrees of freedom at level epsilon can be similarly defined in two dimensions if we do the following. Using the B-spline bases of order k we can define the bilinear spline approximation $S_{k,n_x,n_y}(xy)$ to $f(xy)$ as

$$\hat{f}(xy) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} b_{ij} B_i^{(k)}(x) B_j^{(k)}(y) \equiv S_{k,n_x n_y}(x,y)$$

where the knots have been defined on the Cartesian grid (x_i, y_k) for $i = 1,2,\dots,n_x$, $k=1,2,\dots,n_y$. The degrees of freedom can then be taken as:

$$\text{DoF}(k,\delta,\|\cdot\|_2) = \min\{n_x + n_y \mid S_{k,n_x n_y}(x,y) \text{ is a best } L^2 \text{ approximation to } f \in L^2(0,1)\}.$$

For our preliminary work we obtained the bicubic spline fits from the following equations

$$\hat{f}(x_k, y_k) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} b_{ij} B_i^{(3)}(x_k) B_j^{(3)}(y_\ell)$$

$$k = 1,2,\dots,N \quad \text{and} \quad \ell = 1,2,\dots,N$$

In matrix notation this becomes

$$[f]_{N \times N} = [B_x^{(3)}][b_{ij}][B_y^{(3)}]$$

where $[B_x^{(3)}]$ is an N x n_x matrix, $[b_{ij}]$ is an n_x x n_y matrix and $[B_j^{(3)}]$ is an n_y x N matrix. The solution minimizing the residual, $\sum_{ij} | f(x_i, y_j) - \hat{f}(x_i, y_j)|^2$ is given by

$$[b_{ij}] = \left[[B_x^{(3)}]^T [B_x^{(3)}]\right]^{-1} [B_x^{(3)}]^T [f] [B_y^{(3)}]^T \left[[B_y^{(3)}][B_y^{(3)}]^T\right]^{-1}.$$

For each experiment N was taken to be 128 so that the effects of the quadrature error implicit in this should not be a significant factor, and the residual should be a good approximation to the L^2 error.

For our preliminary work we investigated the possibility of placing the knots for the x and y knot sets from projection of f(xy) along the y and x data sets, respectively. To do this we borrowed an idea suggested by DeBoor <10> wherein he suggests placing the i+1-st knct with respect to the i-th knot for a k-degree spline according to the following

$$\int_{x_i}^{x_{i+1}} \left|f^{(k+1)}(x)\right|^{1/(k+1)} dx = \text{const}$$

where $f^{(i)}(x)$ indicates the i-th derivative of f(x). To determine the x-knot set for a bicubic spline fit the knots were such that

$$\int_{x_i}^{x_{i+1}} \left| \frac{\partial^4}{\partial x^4} \int_{-1}^{1} f(x,y)dy \right|^{\frac{1}{4}} dx$$

and similarly for the y knot set. The results for a bicubic spline fit to $f(xy)=e^{-50(x^2+y^2)}$ with 10 knots determined in this manner for each of the x and y knot sets compared to 10 and 20 knots taken uniformly are shown in figures 1, 2, 3, and 4. The corresponding least squares residuals are tabulated in table 1. For the case of 10 knots the knot placements are shown for DeBoor's algorithm and for uniform knot placement in figures 5 and 6, respectively. It is quite evident in figure 5 that the knots have been concentrated near zero where the function is varying more than at the edges. From the results of table 1 it is clear that 10 knots placed according to the fourth partials of the projections is clearly better than the uniform case for both 10 and 20 knots. Figures 7, 8, and 9 show the results of a bicubic spline approximation to an actual image of an armored personnel carrier, for 40 uniform knots in each direction in figure 7 and for 40

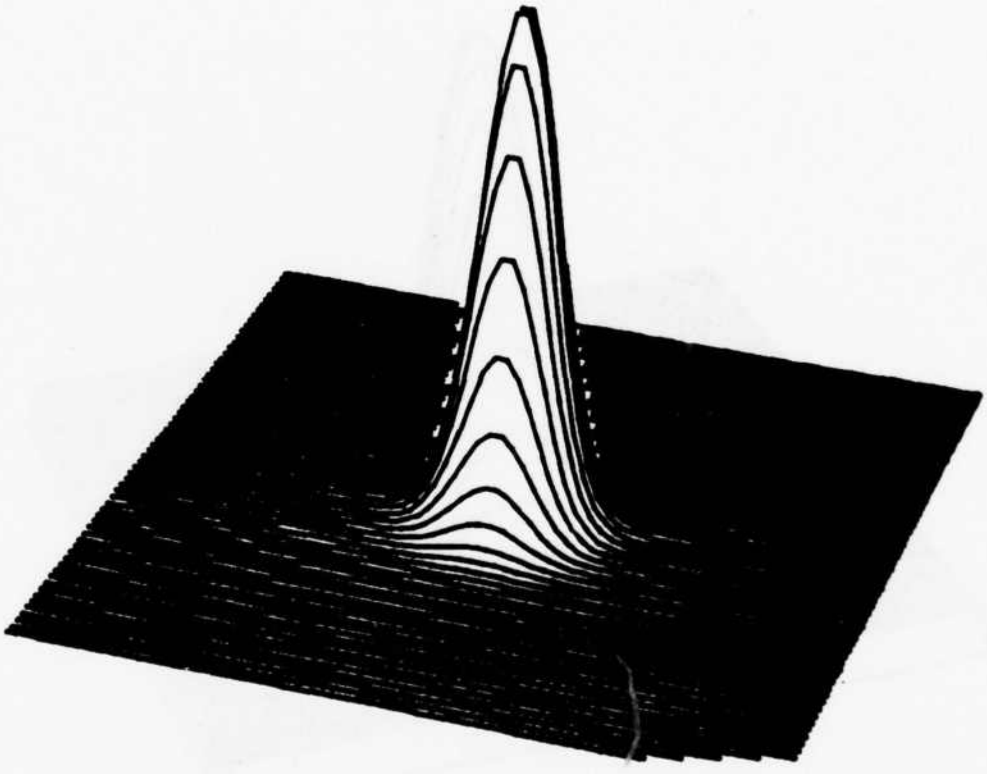


Figure 1. Bicubic Spline Fit with 10 Knots in X and Y Directions Determined from 4th Partials of X and Y Projections

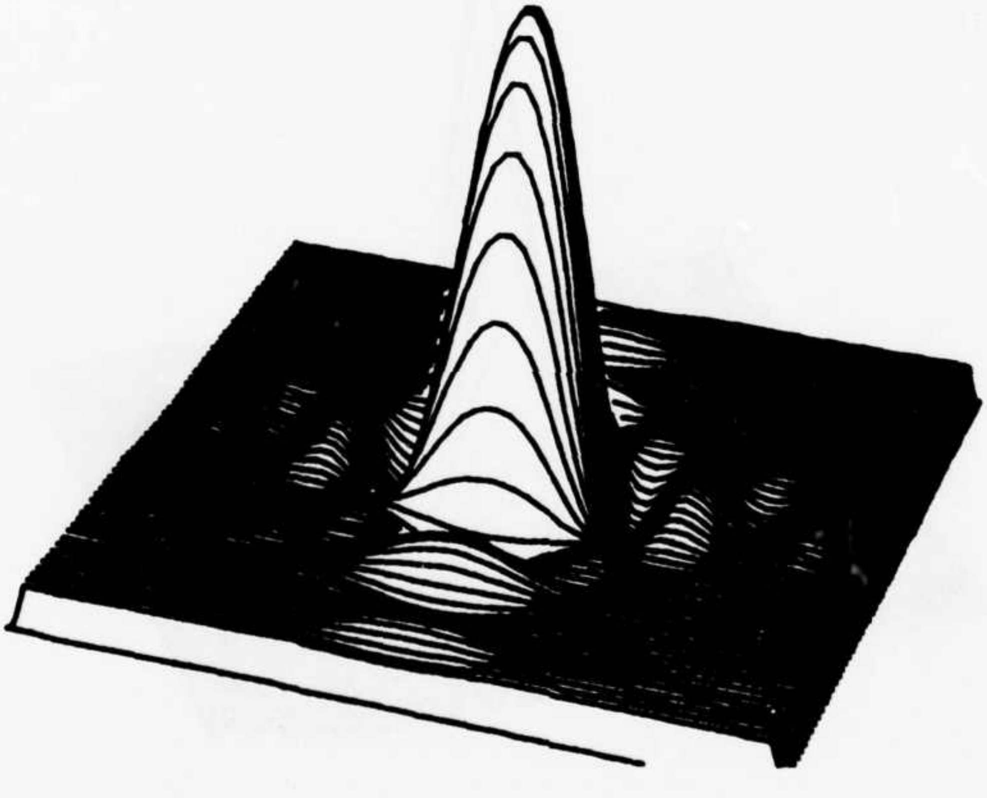


Figure 2. Bicubic Spline Fit with 10 Uniform Knots in X and Y Directions

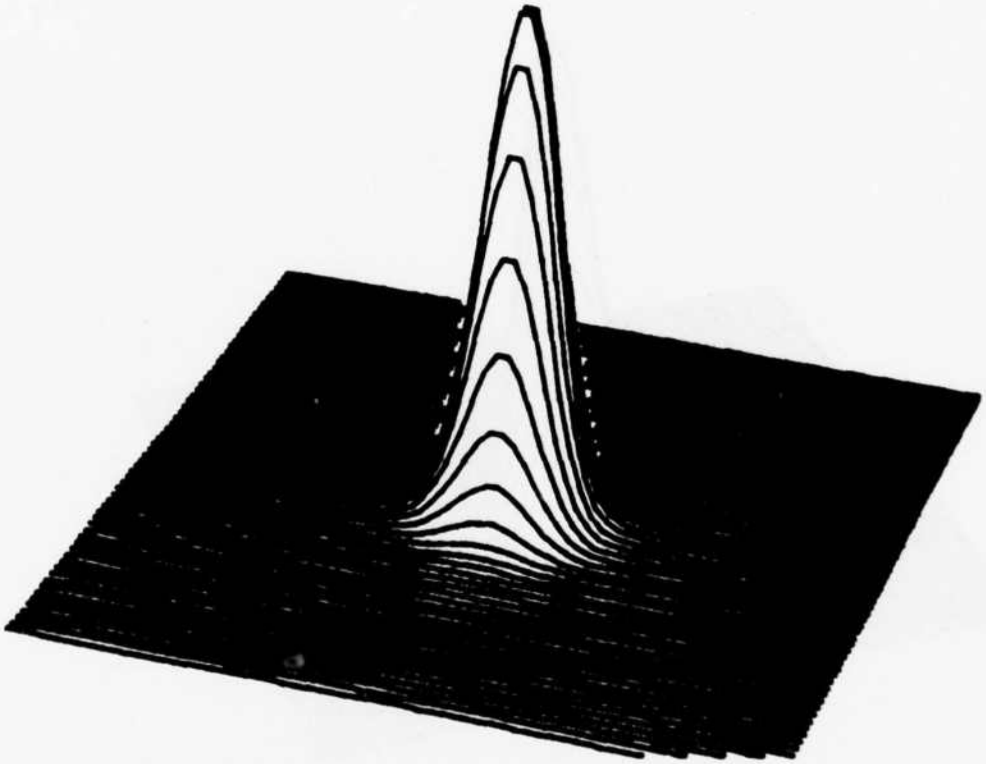


Figure 3. Bicubic Spline Fit with 20 Uniform Knots in X and Y Directions

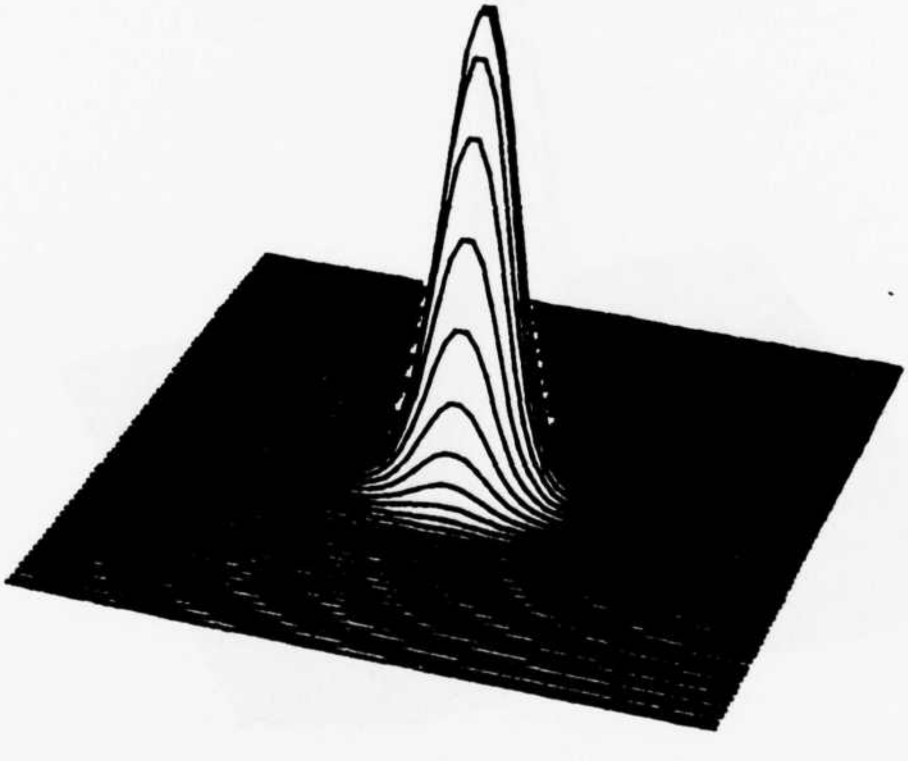


Figure 4. Perspective Plot of Original Function $e^{-50(x^2+y^2)}$

NUMBER OF KNOTS	PLACEMENT MODE	RESIDUAL SUM OF SQUARES
10	UNIFORM	2.158
10	4th DIFFERENCES	1.29×10^{-3}
20	UNIFORM	3.24×10^{-3}

Table 1. Residual Sum of Squares for Knot Placement on Function $e^{-50(x^2+y^2)}$

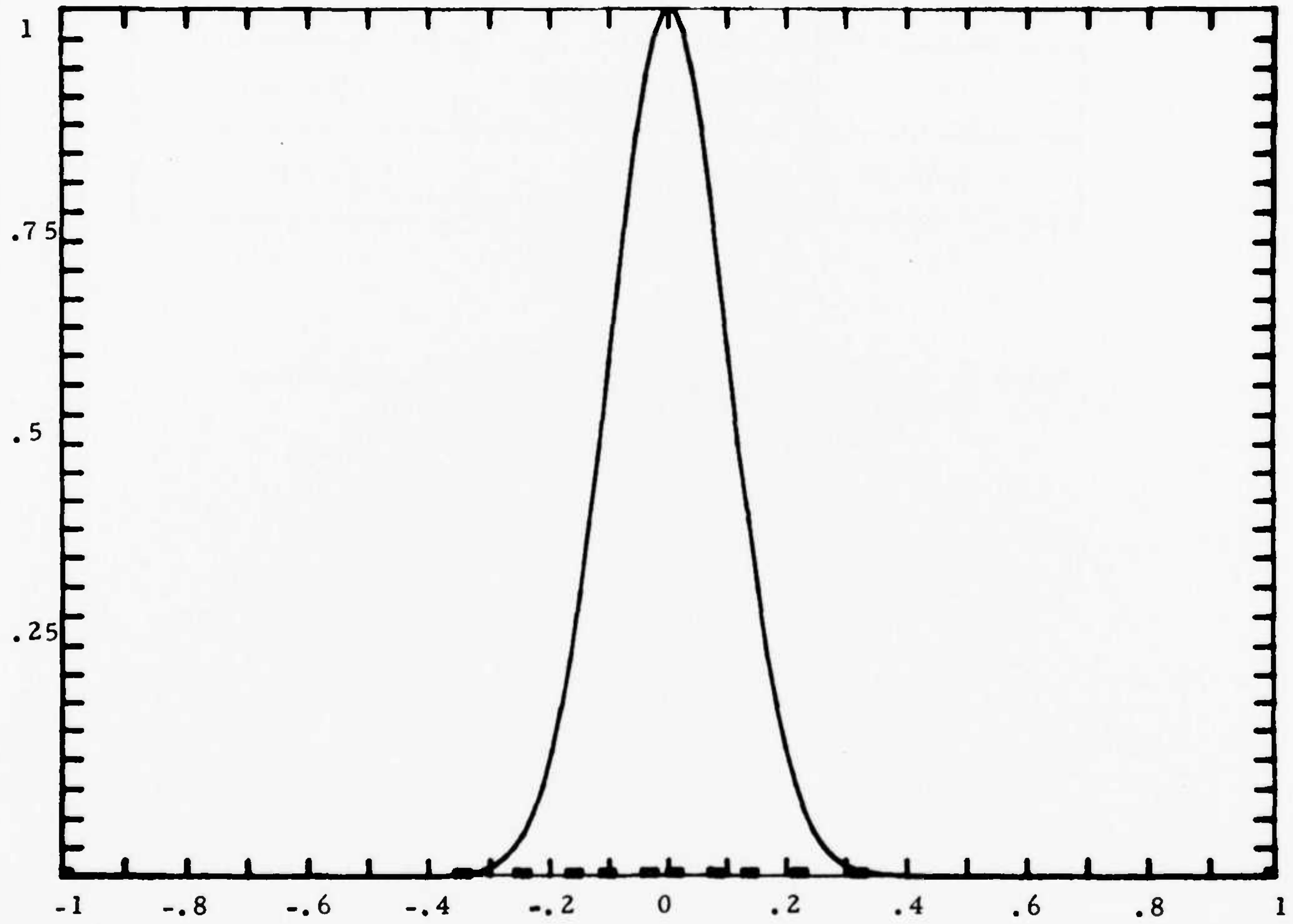


Figure 5. X Projection of f(x, y) and Y-Knot Placement from 4th Partial Algorithm for 10 Knots (Y-Knot Placement Identical)

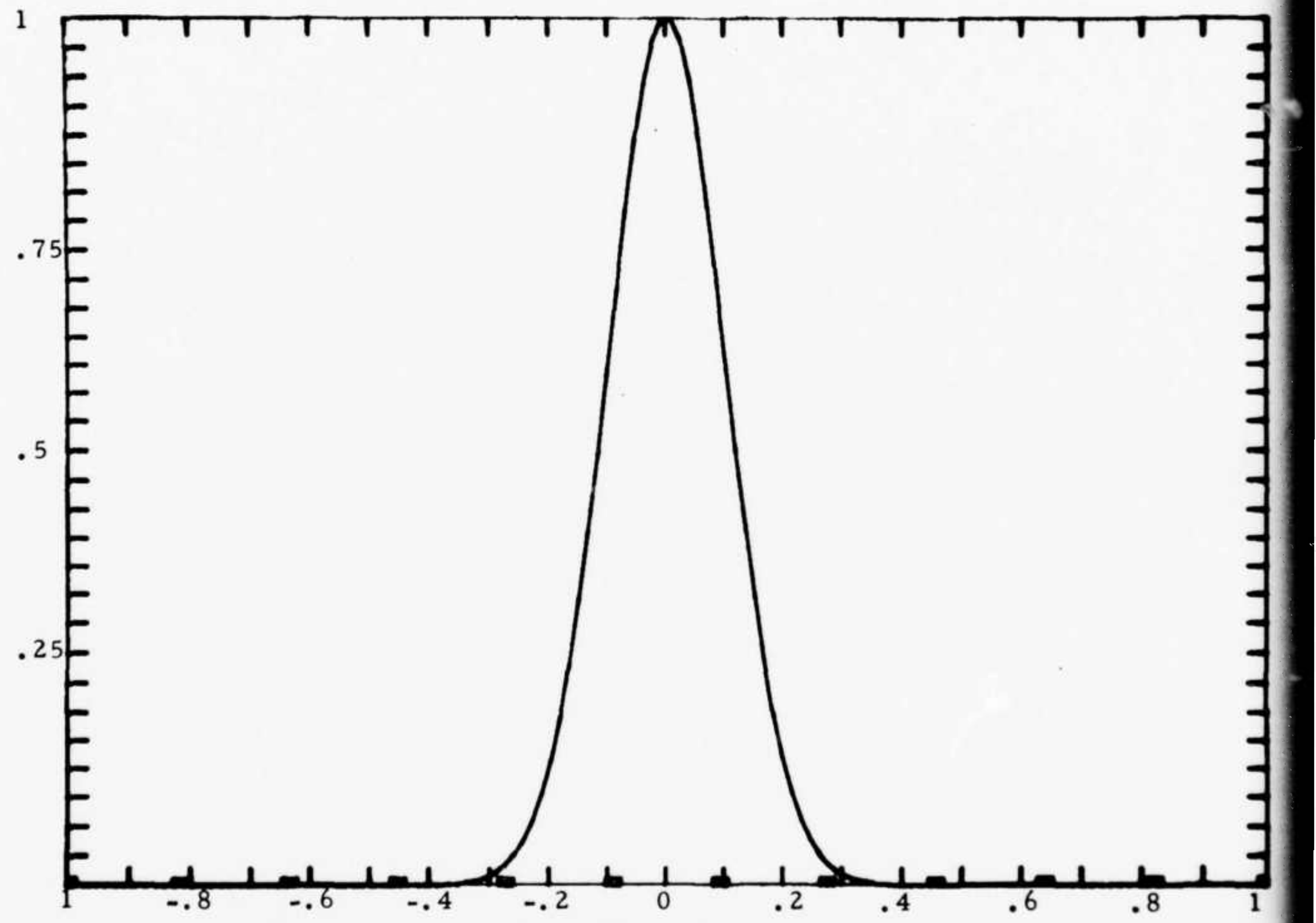


Figure 6. Y Projection and X-Knot Placement for 10 Uniform Knots (Y-Knot Placement Identical)

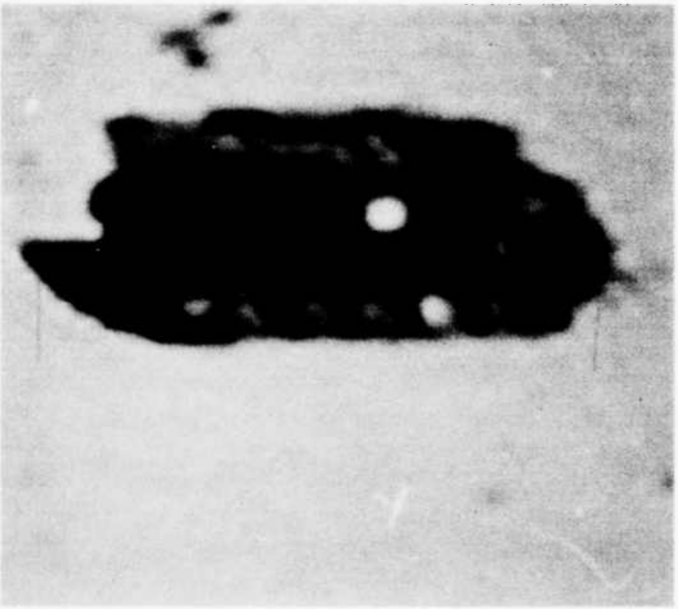


Figure 7. Bicubic Spline Approximation for APC with 40 Knots Uniformly Placed in X and Y Direction

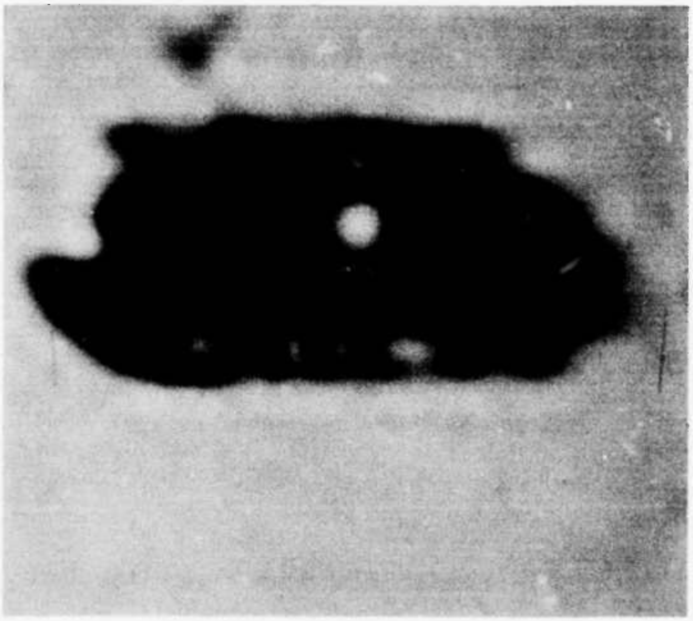


Figure 8. Bicubic Spline Approximation for APC with 40 Knots in X and Y Directions Placed by 4th Differences on X and Y Projections

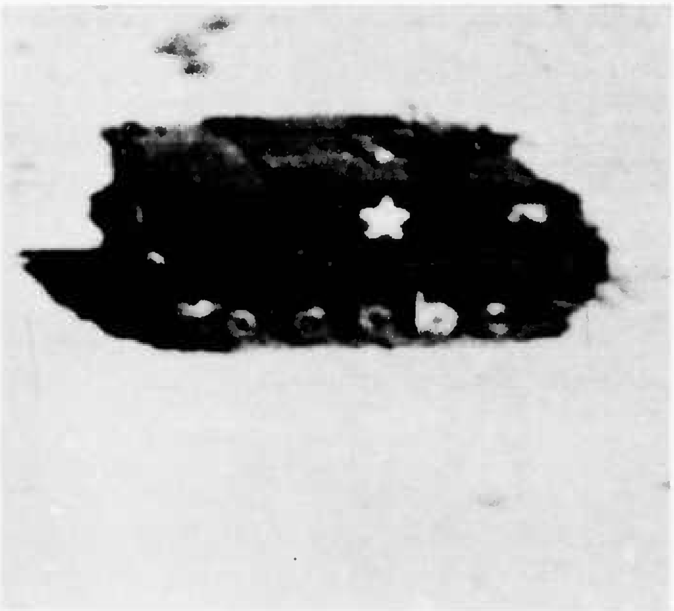


Figure 9. Original APC Image

knots placed according to the fourth difference of the projection data in figure 8. Figure 9 is the original. Table 2 contains the respective residuals. The respective knot placements along with the row and column projections are shown in figures 10 and 11.

Conclusions

In this paper the attempt has been to define the concept of an "epsilon degrees of freedom" of an image and to estimate that degrees of freedom by the number of variable knot bicubic splines necessary to approximate the image of an error level epsilon. Considerable success was achieved for the analytical image $e^{-50(x^2+y^2)}$ where 10 knots, each in the x and y directions were placed by an algorithm dependent on the fourth partials of the projection, and provided an error reduction of three orders of magnitude over the situation where each of the knot sets were uniformly spaced in the x and y directions. The error was such that 10 knots placed in the above manner provided an error lower than that achieved by placing 20 knots uniformly in the x and y directions. For an actual image the results were not so good - mainly a result of the difficulties in numerically obtaining the fourth derivatives. However, while the residual sum of squares is greater for the variable knot case than the uniform knot case more detail is perceptable in the variable knot bicubic spline approximation.

Splines were chosen as the expansion functions owing to their relatively attractive computation and properties and it is felt that with further research an efficient and accurate knot placement algorithm can be developed which would provide better results for real

NUMBER OF KNOTS	PLACEMENT MODE	RESIDUAL SUM OF SQUARES
40	UNIFORM	1.539×10^6
40	4th DIFFERENCES	1.037×10^7

Table 2. Residual Sum of Squares for Knot Placement on APC Image

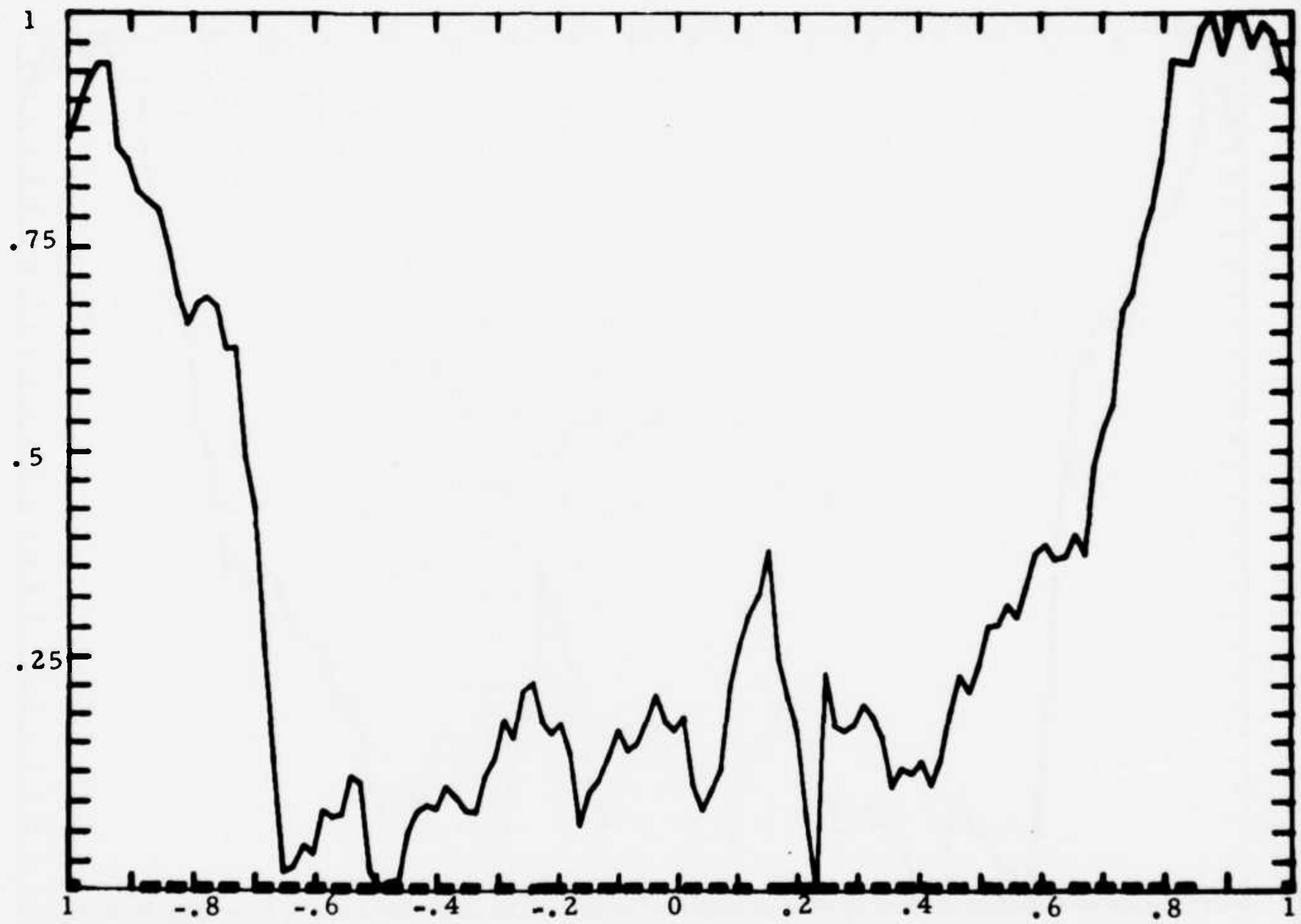


Figure 10a. Knot Placement from 4th Differences on Column Projections, 40 Knots

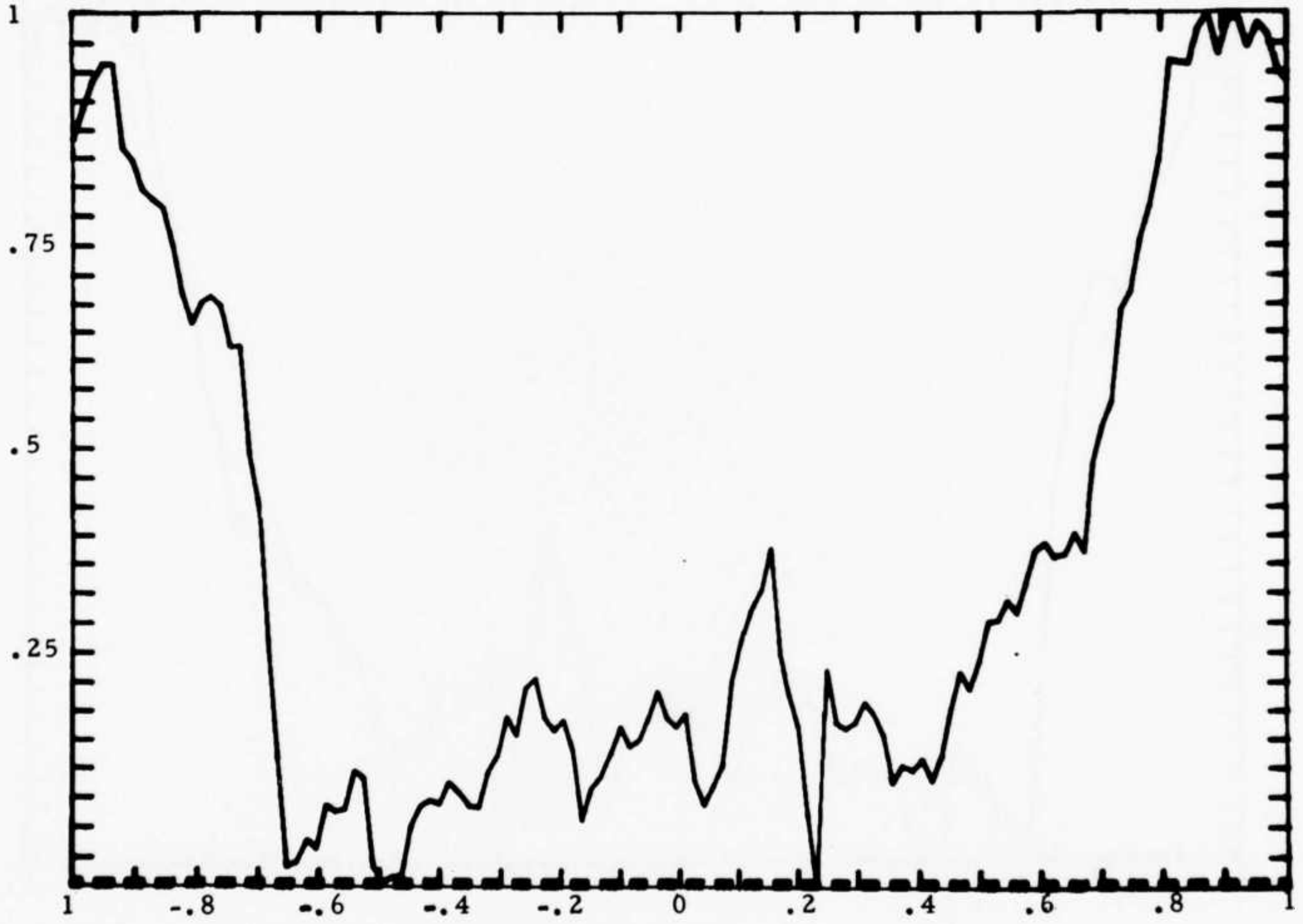


Figure 10b. 40 Knots Placed Uniformly

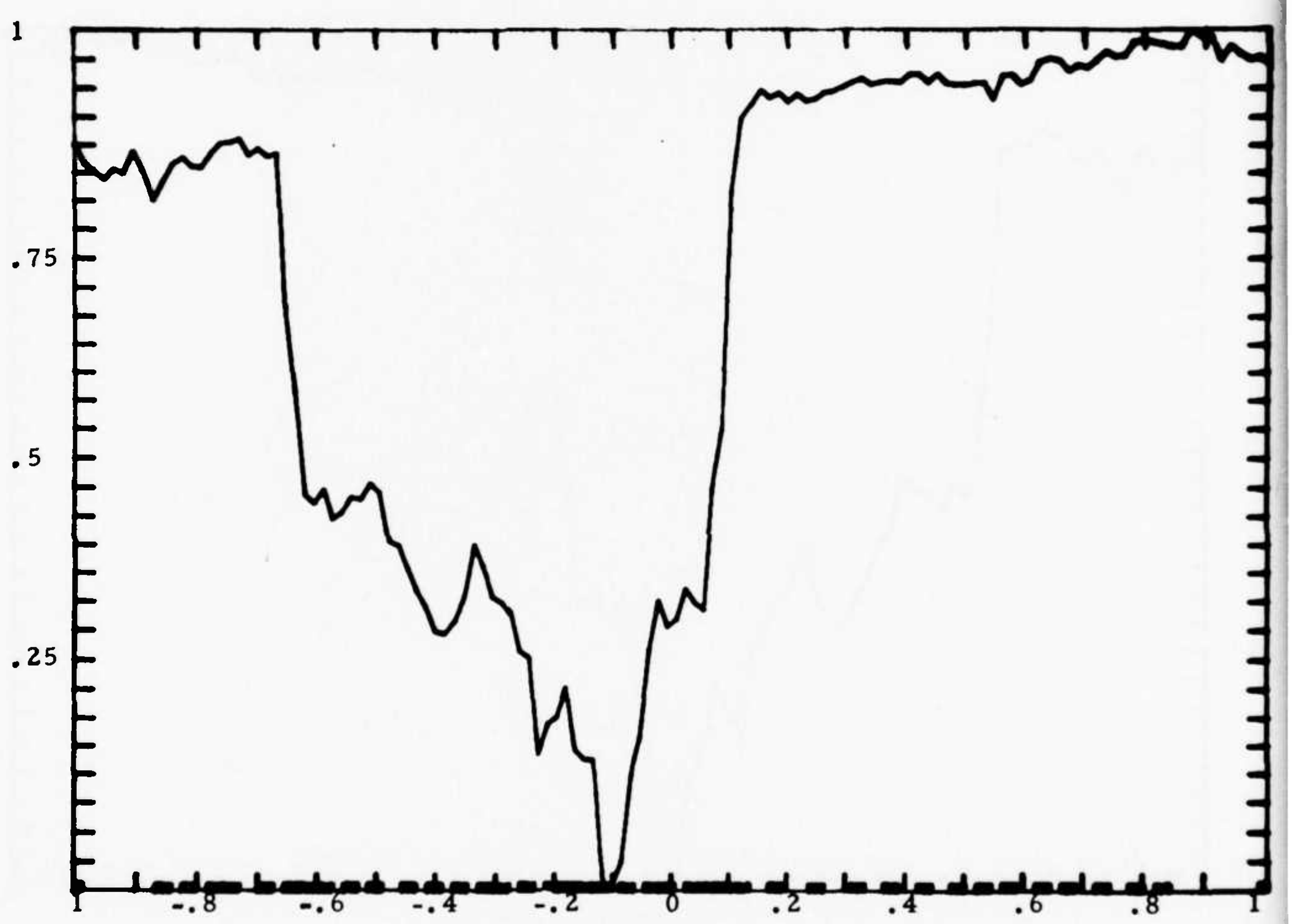


Figure 11a. Knot Placement from 4th Differences on Row Projections, 40 Knots

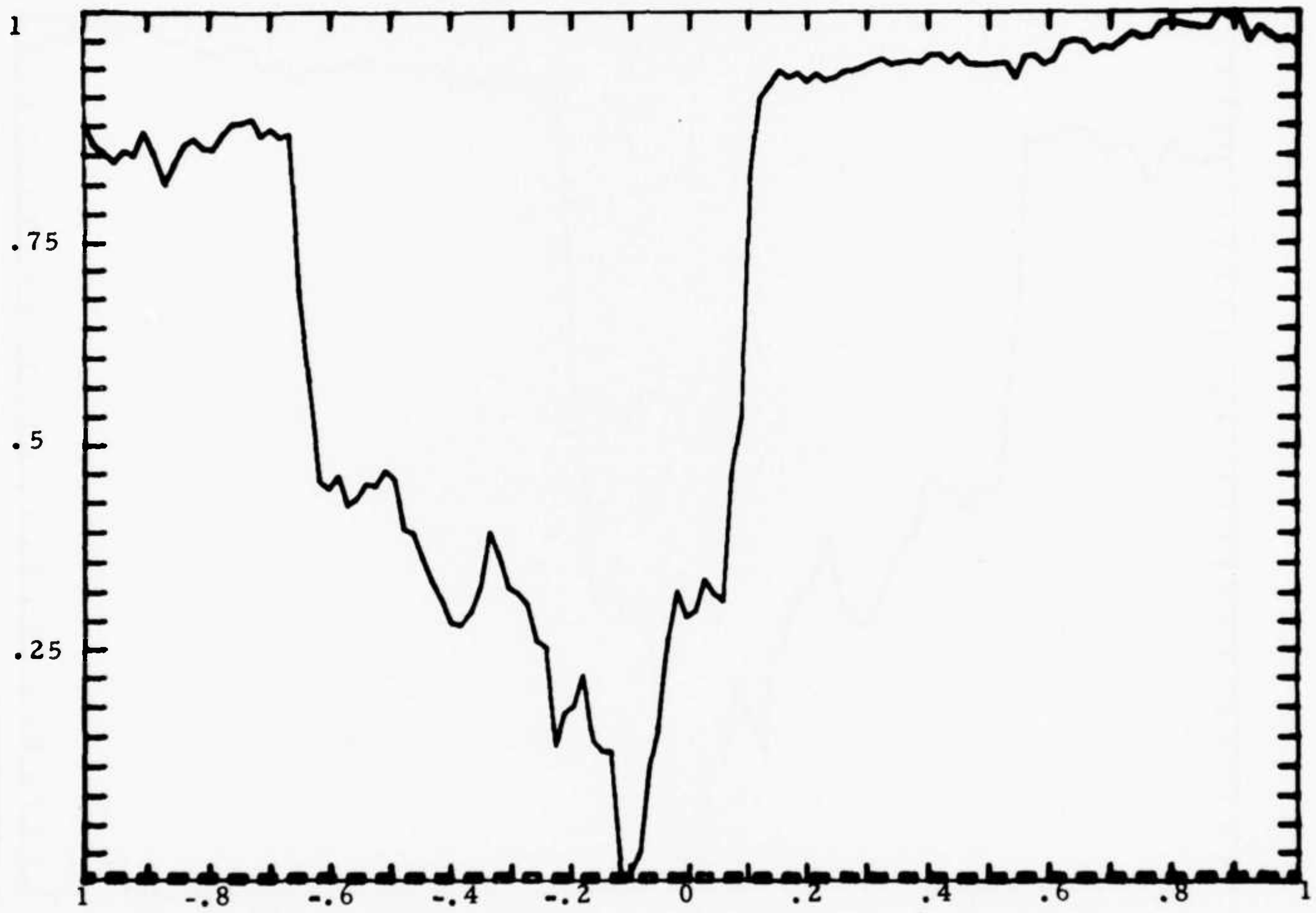


Figure 11b. 40 Knots Placed Uniformly

images than would be obtained for uniform knot bispline approximations.

References

1. S. Twomey, "The Application of Numerical Filtering to the Solution of Integral Equations Encountered in Indirect Sensing Measurements," Journal of the Franklin Institute, Vol. 279, 1965, pp. 95-109.

2. D. McCaughey and H.C. Andrews, "Degrees of Freedom and Projection Imaging," to appear in IEEE Transactions on Acoustics, Speech, and Signal Processing.

3. M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase," IEEE Transactions on Acoustis, Speech, and Signal Processing, Vol. ASSP-24, No. 1, 1976, pp. 58-62.

4. H.C. Andrews and C.L. Patterson, "Singular Value Decompositions and Digital Image Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 1, 1976, pp. 26-53.

5. H.J. Landau and H.O. Pollak, "Prolate Spheroidal Waveforms, Fourier Analysis, and Uncertainty - III: The Dimension of the Space of Essentially Time-and-Band-Limited Signals," Bell Systems Technical Journal, Vol. 41, No. 4, 1962, pp. 1295-1336.

6. D. Slepian and H.O. Pollak, "Prolate Spheroidal Waveforms, Fourier Analysis, and Uncertainty - I," Bell Systems Technical Journal, Vol. 40, No. 1, 1961, pp. 43-64.

7. T. Pavlidis, "Optimal Piecewise Polynomial L Approximations of One and Two Variables," IEEE Transactions on Computers, Vol. C-24, No. 1, 1975, pp. 98-101.

8. C. DeBoor, "On Calculating with B-Splines," Journal of Approximation Theory, Vol. 6, 1972, pp. 50-62.

9. L.L. Schumaker, "Approximation by Splines," from Theory and Applications of Spline Functions, T.N.E. Greville, ed., Academic Press, New York, 1969.

10. C. DeBoor, "Good Approximations by Splines with Variable Knots," from Spline Functions and Approximation Theory, A. Meir and A. Sharma, eds., Birkhausen Verlag Basel und Stuttgart, 1973.

3.2 A TECHNIQUE OF A POSTERIORI RESTORATION

John Morton

Introduction

This project involves the restoration of a blurred image without a priori knowledge of the point spread function (PSF) of the degrading system. That is, consider the model of the degrading system

$$g(x,y) = \int\int_{-\infty}^{+} h(h,y,\alpha,\beta)\ f(\alpha,\beta)d\alpha d\beta + n(x,y)$$

where $g(x,y)$ is the degraded image, $h(x,y,\alpha,\beta)$ is the system PSF, $f(x,y)$ is the image, and $n(x,y)$ is noise.

The majority of restoration techniques assume knowledge of $h(x,y,\alpha,\beta)$ and estimate $f(x,y)$. Yet, in most cases one would not have knowledge of $h(x,y,\alpha,\beta)$. Thus, the technique under study estimates both $h(x,y,\alpha,\beta)$ and $f(x,y)$ solely from $g(x,y)$ while assuming the following:

1) the PSF is spatially invariant, i.e., $h(x,y,\alpha,\beta) = h(x-\alpha,y-\beta)$,
2) the extent of the PSF is small compared to the extent of the image, and
3) the image is not so severely blurred such that one cannot tell the general class, for example, building, outdoor scene, etc., to which the blurred image belongs.

Consider the Fourier transform of h

$$F\{h\} = H(u,v)$$

Note in general that H is complex and can be represented in magnitude phase form as

$$H(u,v) = |H(u,v)|\, e^{j\theta(u,v)}$$

Thus, if one knows $|H(u,v)|$ and $\theta(u,v)$, one in effect knows $h(x,y)$ via the inverse Fourier transform relationship. Previous work in this area was termed "blind deconvolution" <1-4> and has had some moderate successes. The work of Cole <1> estimated $|H(u,v)|$. Cannon <2-4> extended this work by estimating $|H(u,v)|$ and estimating $\theta(u,v)$ for a set of three common blurs. The estimate of $\theta(u,v)$ was accomplished by assuming the blur was one of three possible blurs. Using a pattern recognition technique Cannon could discriminate between the three possible blurs and effectively estimate $\theta(u,v)$. The present work does not assume a finite set of possible blurs but will estimate $\theta(u,v)$ directly from $g(x,y)$. The mathematics involved in the estimate of $\theta(u,v)$ is similar to the mathematics developed by Knox [5-6] for recovering astronomical images degraded by turbulance.

Schematically, the a posteriori restcration is presented in figure 1.

Estimate of PSF Magnitude

The estimate of $|H(u,v)|$ will be essentially accomplished by the techniques developed by Cole <1> and Cannon <2>. For completeness, however, let us review the methods of estimation. For reasons of clarity and due to the progress of the work to date, only the noise-free case will be considered herein. By assumption

$$g(x,y) = \int\int_{-\infty}^{\infty} h(x-\alpha,\ y-\beta)\ f(x,y)$$

$$= h(x,y) * f(x,y)$$

where * denotes convolution.

If the blurred image $g(x,y)$ is broken into smaller images $g_i(x,y)$ such that the spatial extent of $h(x,y)$ is less than the spatial extent of each $g_i(x,y)$, then approximately

$$g_i(x,y) \approx h(x,y) * f_i(x,y)$$

In addition

$$G_i(u,v) \approx H(u,v)F_i(u,v) \qquad (1)$$

where upper case letters denote Fourier transforms and specifically

$$|G_i(u,v)| \approx |H(u,v)|\,|F_i(u,v)|.$$

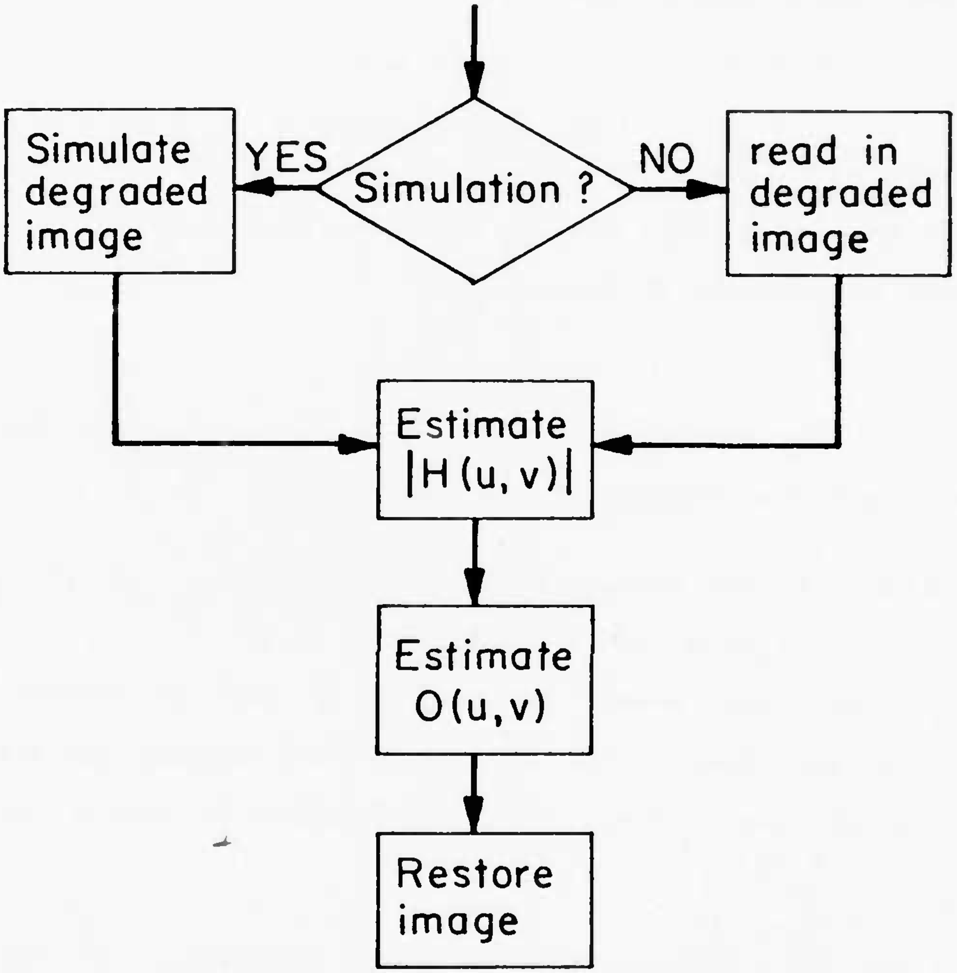


Figure 1. A posteriori restoration schematic.

Averaging over the ensemble of subimages where <·> denotes averaging over quantities indexed by i,

$$<|G_i(u,v)|> \approx < |H(u,v)| \, |F_i(u,v)| >$$

$$\approx |H(u,v)| < |F_i(u,v)| >.$$

For classes of images, for example, buildings, it was found <1> that $<|F_i(u,v)|>$ was very similar for different images of the given class. Thus, an estimate of $|H(u,v)|$ is

$$|H(u,v) \approx \frac{<|G_i(u,v)|>}{<|F_i^P(u,v)|>}$$

where $<|F_i^P(u,v)|>$ is performed on any image belonging to the same class or averaged over members of the same class.

Alternately, if one assumes stationarity,

$$\Phi_g(u,v) = |H(u,v)|^2 \, \Phi_F(u,v)$$

where Φ_g, Φ_f are the power spectra of g and f, respectively. Similarly, it was found that $\Phi_f(u,v)$ is very similar for different images of a given class. Thus, a second estimate of $H(u,v)$ is

$$|H(u,v)| \approx \left[\frac{\Phi_g(u,v)}{\Phi_F^P(u,v)}\right]^{\frac{1}{2}}$$

where Φ_f^P is the power spectrum of an image belonging to the same class.

Estimation of PSF Phase

The estimate of $\theta(u,v)$ proceeds as follows.

$$j\phi_i(u,v) \quad j\theta(u,v) \quad j\varphi_i(u,v)$$

Let us consider the autocorrelation

$$R_G(u,v,\Delta u,\Delta v) = <G_i(u,v) \, G_i^*(u+\Delta u, v+\Delta v)> \qquad (2)$$

Substituting G_i (u,v) from equation (1) into equation (2)

$$R_G(u,v,\Delta u,\Delta v) \approx < H(u,v)F_i(u,v)\, H^*(u+\Delta u, v+\Delta v)\, F_i^*(u+\Delta u, v+\Delta v)>$$

$$= H(u,v)\, H^*(u+\Delta u, v+\Delta v) < F_i(u,v)F_i^*(u+\Delta u, v+\Delta v) > \qquad (3)$$

$$= H(u,v)\, H^*(u+\Delta u, v+\Delta v)\, R_F(u,v,\Delta u,\Delta v);[\theta(u,v)-\theta(u+\Delta u, v+\Delta v)]$$

$$= |H(u,v)|\, |H^*(u+\Delta u, v+\Delta v)|\, e \qquad R_F(u,v,\Delta u,\Delta v)$$

From equation (1)

$$|H(u,v)|\, |H^*(u+\Delta u, v+\Delta v)| \approx \frac{|R_G(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)|} \,. \qquad (4)$$

Combining equations (3) and (4) and rearranging results,

$$e^{;[\theta(u,v)-\theta(u+\Delta u, v+\Delta v)]} \approx \frac{R_G(u,v,\Delta u,\Delta v)}{|R_G(u,v,\Delta u,\Delta v)|} \; \frac{|R_F(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)}$$

and it follows that

$$\theta(u+\Delta u, v+\Delta v) \approx \theta(u,v) - \tan^{-1}\left\{ \frac{R_G(u,v,\Delta u,\Delta v)}{|R_G(u,v,\Delta u,\Delta v)|} \quad \frac{|R_F(u,v,\Delta u,\Delta v)|}{R_F(u,v,\Delta u,\Delta v)} \right\}$$

Assuming one knows R_G and R_F and letting θ(0,0) = 0, one may obtain θ(u,v) by the above recursive relationship. Note, however, that one can calculate R from the degraded image but one does not know R a priori. Thus to estimate (u,v) one must estimate R . It has been assumed that the image is not so severely degraded that one cannot tell the general class to which it belongs. As a result, the estimate of R will be calculated by using a similar clear image. Thus, the success of the phase estimation technique under study is contingent on the similarity of R for different images within a given class.

Progress to Date

A series of computer programs which estimate both magnitude and phase of the PSF has been written. The programs are being run on test cases to ascertain whether or not the programs are error-free. The programs have not been tried on real images. In particular the important assumption of similarity of autocorrelation R for different images within the same class has not as yet been experimentally verified. This check will proceed as soon as reasonable certainty as to the correctness of the computer programs has been determined.

Although the tests to date are somewhat artificial, they reveal problem areas and illustrate the difficulty of obtaining a good check of the computer programs. For example, consider the following test case. The function

$$f(x,y) = 200\, e^{-8|x| - 12|y|}$$

was calculated and stored on a file in standard eight bit/pixel image format. This simulated image was then convolved with a function h(x,y) and the result stored on another file in standard eight bit/pixel image format. Because the object of the test was to test the computer programs, the estimate of H(u,v) was then calculated using the known f(x,y) and as a consequence R was not estimated but was calculated from the known f(x,y).

Results along the u = 0 and v = 0 axis are shown in figures 2 and 3, respectively. The correct answers are

$$H(u,0) = \frac{\sin(5\pi u)}{5\pi u}$$

and

$$H(0,v) = \frac{\sin(5\pi v)}{5\pi v}$$

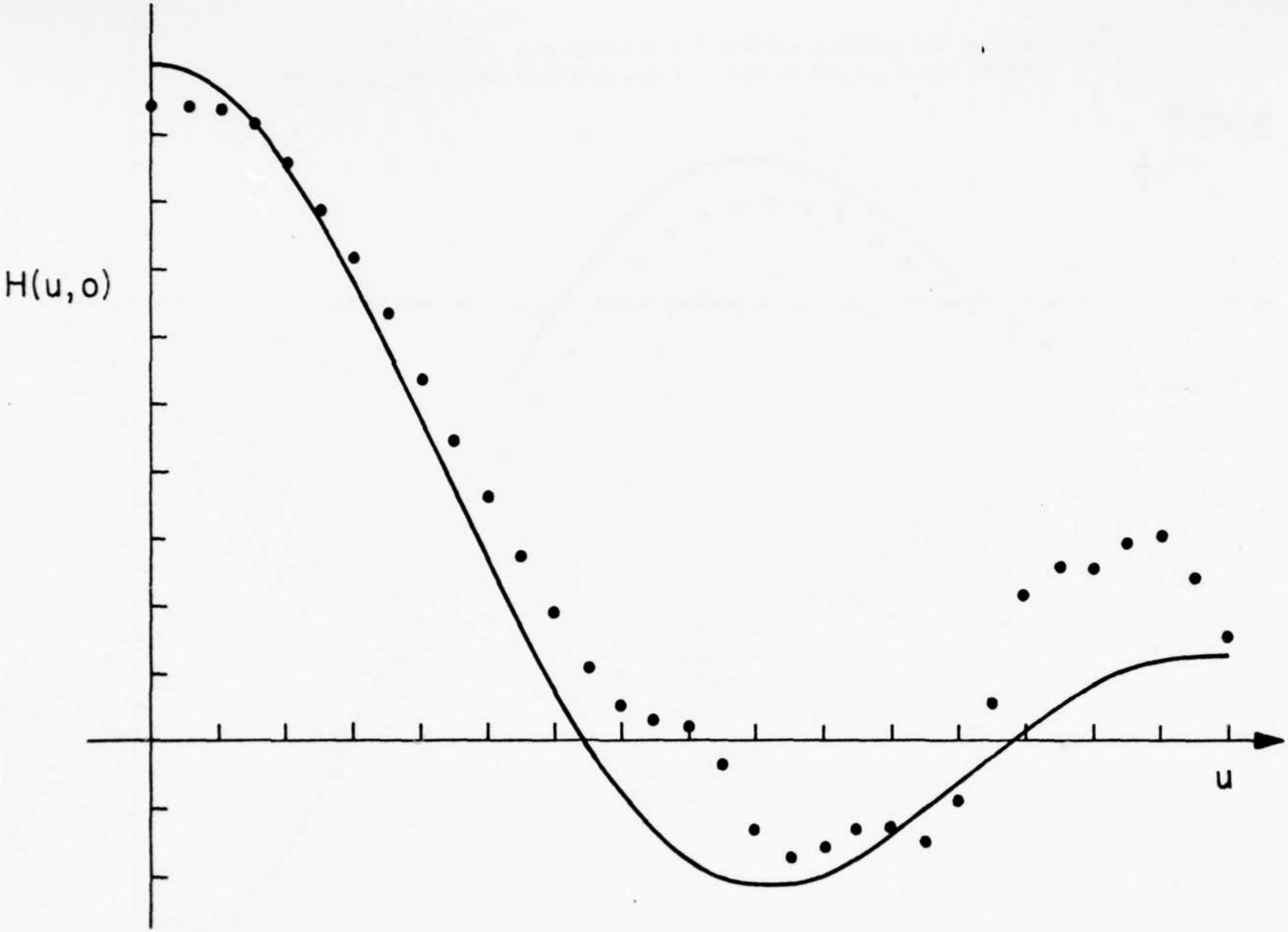


Figure 2. Computer program check. The solid line is the known answer. The dots are the results obtained via the computer programs.

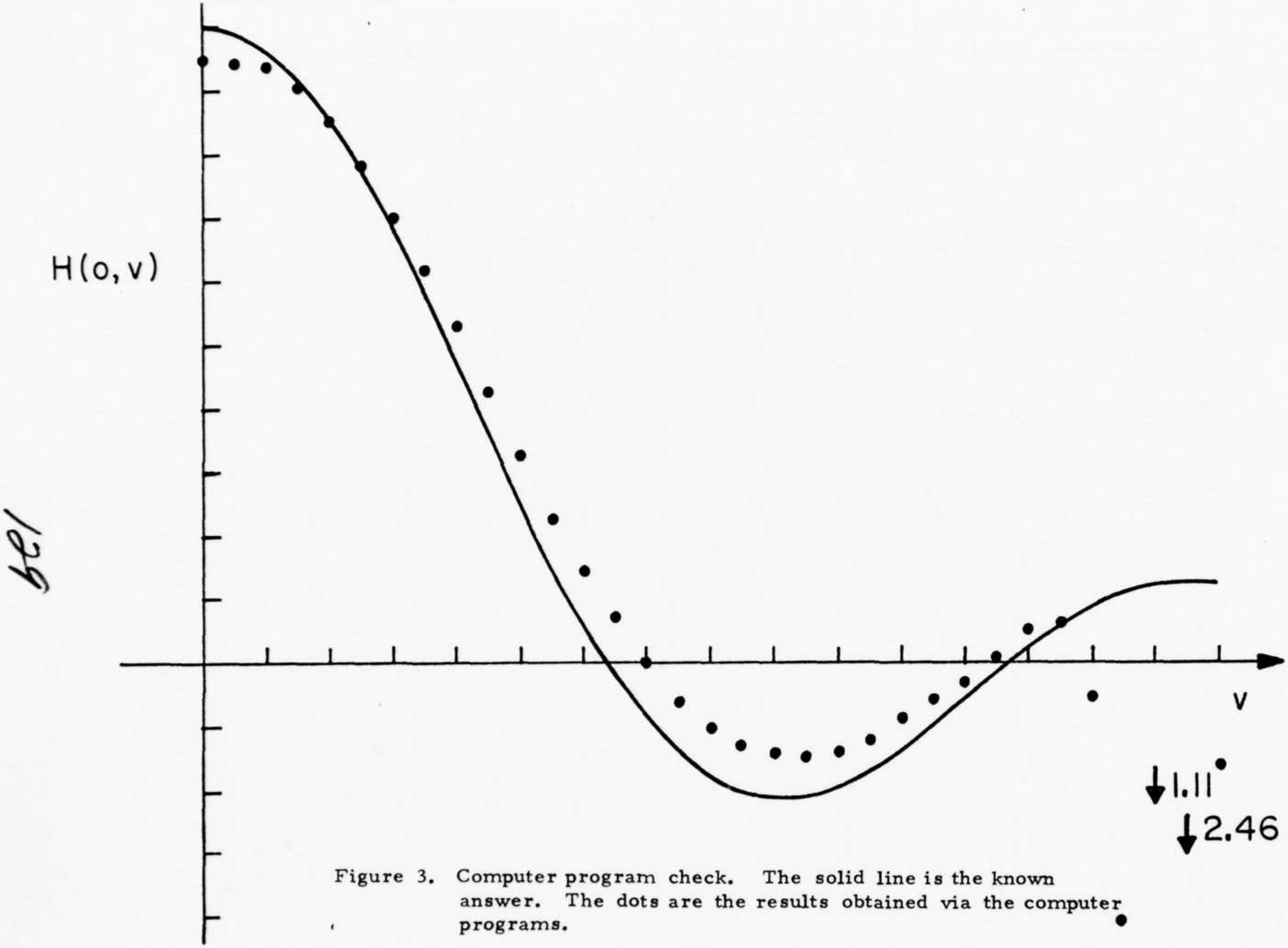


Figure 3. Computer program check. The solid line is the known answer. The dots are the results obtained via the computer programs.

respectively. Note in both figures 2 and 3 that the estimated values are approximately correct and in general the error increases for increasing u and v. It is believed that the error is due to the estimation procedure and not the computer programs. Sources of error include the discretization of a continuous process, limitations inherent in using eight bit/pixel images, and inaccuracies in using the discrete Fourier transform to approximate the continuous Fourier transform.

References

1. E.R. Cole, "The Removal of Unknown Image Blurs by Hcmomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSc-74-029, June 1973.

2. T.M. Cannon, "Digital Image Deblurring by Ncnlinear Homomorphic Filtering," Department of Computer Science, University of Utah, ARPA Technical Report UTEC-CSc-74-091, August 1974.

3. T.G. Stockham, T.M. Cannon, and R.B. Ingebritsen, "Blind Deconvolution through Digital Signal Processing," Proceedings IEEE, Vol. 63, No. 4, April 1975, pp. 678-692.

4. T.M. Cannon, "Blind Deconvolution of Spatially Invariant Image Blurs with Phase," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 1, February 1976, pp. 58-63.

5. K.T. Knox and B.J. Thompson, "Recovery of Images from Atmospherically Degraded Short-Exposure Photographs," The

Astrophysical Journal, Vol. 193, October 1974, pp. 245-248.

6. K.T. Knox, "Diffraction-Limited Imaging with Astronomical Telescopes," Ph.D. Dissertation, University of Rochester, Rochester, New York, 1975.

3.3 IMAGE PROCESSING BY TRANSFORMS OVER A FINITE FIELD USING THE OVERLAP-ADD SECTIONING TECHNIQUE

Jin Soh

The fast Fourier transform (FFT) has been used to reduce the computational cost of convolution. Recent work has demonstrated that there are attractive analogies to the FFT in certain finite fields and rings. Transforms using these number theoretic concepts have been developed to compute convolutions without loss of information of sequences using the FFT algorithm. The transforms are defined on finite fields and rings of integers with the arithmetic carried out modulo an integer. Because of these characteristics they are ideally suited for digital computation by taking into account quantization of amplitude as well as time in their definitions. When the modulus is chosen as a Fermat number, transformation requires only on the order of N log N additions and word shifts but no multiplications. In addition to being efficient, they have no roundoff errors and do not require storage of basis functions. There is a restriction on the length of sequences imposed by the word length and also there is a problem of overflow but methods for overcoming these are being investigated. Comparison with FFT has shown a substantial improvement in efficiency and accuracy. Variations on the basic number theoretic transforms are also being developed.

It is the purpose of the on-going experiment to point out that the two-dimensional number theoretic transform may be defined as a simple extension of the one-dimensional number theoretic transform. For practical cases, two-dimensional finite impulse response filters involve impulse response of a lot fewer than ordinary picture size 256 x 256 or 512 x 512 in almost all cases. Since these number theoretic transforms have the difficulty that the word length required is proportional to the length of the sequence being convolved, it is conceivable that sectioning techniques may be of value when sequences must be convolved exactly, such as when multiplication is very costly, or when special-purpose hardware for modest length convolution is being considered. For the large picture mentioned above the convolution can be done by sectioning the picture by efficient overlap-add sectioning deleting and end effect. For y(n) = x(n) * h(n), we assume that x(n) is of long duration or that the number of samples representing x(n) exceeds the size of the h(n). Then, since the shift of the convolution y(n) is simply the sum of the shifts of the functions being convolved, no information is lost if we shift each sectioned function to the origin prior to convolution. As a result, it is necessary to decompose x(n) into sections and compute the discrete convolution as many smaller convclutions to avoid the problem of dynamic range of the number theoretic transform. At the same time, if N is small, then we would have the advantage that considerable time saving could be achieved by using shorter sequences. Therefore it can be true that the use of number theoretic transform as a convolution tool may make it possible to have faster and more efficient picture processing.

References

1. R.C Agarwal and C.S. Burrus, "Number Theoretic Transforms to Implement Fast Digital Convolutions," IEEE Proceedings, Vol. 63, No. 4, April 1975.

2. I.S. Reed and T.K. Truong, "The Use of Finite Fields to Compute Convolutions," IEEE Transactions on Information Theory, Vol. IT-21, No. 2, March 1975.

3. J.M. Pollard, "The Fast Fourier Transform in a Finite Field," Mathematics of Computation, Vol. 25, No. 114, April 1971.

4. E.O. Brigham, The Fast Fourier Transform, Prentice-Hall, Inc., 1974.

3.4 PHASE CODING FOR OPTICAL IMAGE PROCESSING

Chung-Kai Hseuh and Alexander A. Sawchuk

Many operations in digital image processing such as restoration, linear filtering and interpolation can be performed with analog optical systems. Research in these fields in the past has generally concentrated on coherent optical systems as opposed to incoherent systems. With the use of general amplitude and phase optical elements, many useful image processing operations can be done very easily with analog incoherent systems. If general optical elements combining amplitude and phase variation can be made, the only major restriction on incoherent systems is that the impulse response must be non-negative.

A very useful function suitable for an incoherent optical processing system is a B-spline interpolator to give a continuous desampled output from the discrete spots on a CRT or other display system. Other useful operations include Wiener filters, high spatial frequency emphasis filters, edge or directionally sensitive filters, or restoration filters.

The filtering operation of the optical system can be described by a complex amplitude point-spread function $\underline{h}(x)$ or non-negative intensity response equal to the magnitude-squared $|\underline{h}(x)|^2$. In the frequency or pupil domain, omitting scaling factors for simplicity, the desired complex pupil function $\underline{H}(f)$ is given by

$$\underline{H}(f) = \mathcal{F}\{\underline{h}(x)\} \qquad (1)$$

where $\mathcal{F}\{\cdot\}$ denotes the Pourier transform operation. The incoherent optical transfer function $\mathcal{H}(f)$ is then described by

$$\underline{\mathcal{H}}(f) = \underline{H}(f) * \underline{H}(f) = \mathcal{F}\{|\underline{h}(x)|^2\} \qquad (2)$$

using transform theory.

With these definitions, the problems and tradeoffs of optical processing can be defined. Given a desired $\underline{h}(x)$ or $|\underline{h}(x)|^2$, it has been difficult to directly synthesize anything other than simple functions $\underline{H}(f)$. With the development of some new techniques in complex computer hologram synthesis, it is hoped that this fundamental limitation can be overcome. One major approach that will be investigated is to use the precision plotting capabilities of a computer controlled microdensitometer to plot amplitude/phase masks directly. Even with improved capabilities of these areas, the problems of spatial frequency response and dynamic range of the hologram recording medium still exist, since desirable functions $|\underline{h}(x)|^2$ generally have transforms which reach some of the extremes. Here, there is fortunately a free variable which can be manipulated to alleviate some of these problems. This variable is the phase $<$ h(x) of the impulse response which is entirely free. Thus, the major problem of incoherent processing is to choose a pupil function H(f) which has low dynamic range in amplitude and phase, has low spatial frequency structure, making it easy to plot, and gives the desired intensity response h(x) accurately. This type of problem is related to other problems in image processing and is referred to under the general name of "phase coding."

One possible approach is to obtain a constant amplitude $|H(f)|$ for the pupil filter. This eliminates the need for any amplitude variations. In digital holography the resulting pupil filter is called a kinoform. One approach to get the constant amplitude spectrum in the transform domain is to assume a phase diffuser, either random or deterministic, for the phase of h(x). Another approach is to start with a random phase in the object domain and go through a recursive searching routine to reduce the dynamic range, complexity, and object error. Although this numerical searching is a time-consuming process in general, it does show better results than using a diffuser alone. In the application of analog filtering, the hologram is computed only once and the computation time is no longer a serious problem as long as good impulse response has been plotted.

One phase coding method was first developed at IBM <1> for the coding of kinoforms. Kinoform errors due to truncation and quantization are discussed by Gallagher and Liu <2>. In this method, the phases of object are assigned randomly and the object is then Fourier transformed. The transform is set to have a constant amplitude and is then inversely transformed. The resultant phase is used as the new phase for the object. This procedure is repeated until an error criterion is met.

Although this method generally gives good results, it has been found <3> that the error at some points fails to decrease, stabilizing at a high level. This prevents the error from being arbitrarily small. To overcome this deficiency, Fienup <3> adopts an Input-Output Approach, which allows the amplitude of the input as well as the phase to be modified. This gives better control of the reconstruction. In this model if certain output is used as input then the same output is obtained. Hence modification can be either added to the previous input or present output. We will refer to these two methods as Fienup I and Fienup II, respectively.

In general, the shape of the spectrum does not have to be flat. The problem then becomes: given the amplitudes of the object and transform, how do we find the corresponding phases? This problem has been studied by Gerchberg and Saxton <4> and the method is essentially the same as the iterative method mentioned above except the amplitude of the transform is set to be the desired spectrum rather than a constant. Convergence of this method is also proved. This method has found applications in many fields such as microscopy and stellar interferometry.

The last method suggests a modification to the kinoform procedure. We can allow the amplitude of the transform to vary in such a way that the amplitude is still smooth enough to plot. Better results and faster convergence could be expected. Since the spectrum can always be expanded in a Fourier series, the sum of the first few large components can be chosen as the approximation of the spectrum. Here we can consider the first method <1> as a special case in which only D.C. components are chosen. Note that Fienup's method cannot be applied in this general approach since Fienup's method is based on the assumption that the amplitude of the transform is constant. Further derivation is required in applying this method.

In evaluating the results, we need a good error criterion which can be used to modify the computing procedure or to terminate the procedure without human intervention. This error criterion has to be consistent with the human visual system. In general, this task is not possible and instead the mean squared amplitude error <2> or the root-mean squared intensity error <4> is used. Since the actual quantity measured is the intensity of the image, the latter one might be a better criterion although it is larger in general. We can also argue that since human eyes are sensitive to relative intensity other than absolute intensity, it would be better tc normalize the above quantities so that these quantities are not affected by the absolute intensity and, hopefully, the specific object used. The error criterion adopted in this study is the following for an N x N picture

$$E = \left[\frac{1}{N^2} \sum_{M,N=0}^{N-1} \left(|a_{mn}|^2 - \tau_{mn}^2 \right)^2 \right]^{\frac{1}{2}} / \tau_{max}^2 \qquad (3)$$

where $\hat{a}_{mn}$, τ_{mn} are the resultant image and the desired image, respectively and

$$\tau_{max} = \underset{m,n}{\mathrm{Max}}\ \tau_{mn} \quad . \qquad (4)$$

One dimensional objects are used in the experiment for simplicity. Only one of them will be illustrated here. Figure 1 shows one line of a picture. This is a 256 point array imbedded in a 512 points array in order to smooth the transform and thus the kinoform. Figures 2, 3, and 4 show the resultant images and kinoforms after 34 iterations by the methods of Hirsch <1>, Fienup I and Fienup II <4>. Compared with figure 1, we found that Fienup II is slightly better but almost the same as the first method. Fienup I, however, gets nearly perfect reconstruction in the region of the object while more noise is introduced elsewhere. Since intensity is measured in practice, this noise in amplitude contributes less in intensity.

The corresponding errors evaluated by eq.(1) are given by

$$E_1 = .05222 \quad E_2 = .01297 \quad E_3 = .04449$$

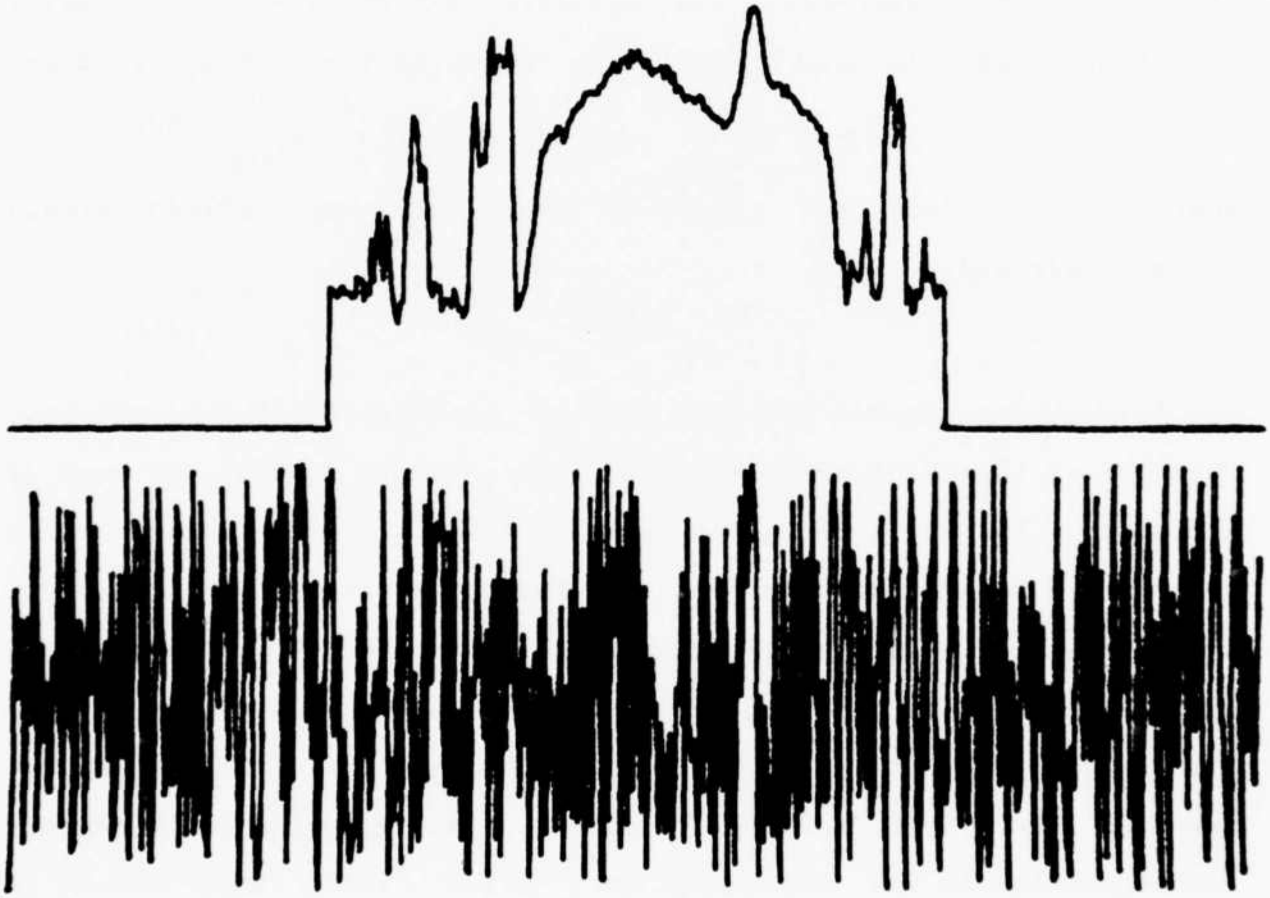


Figure 1. One line of a picture with initial random phase

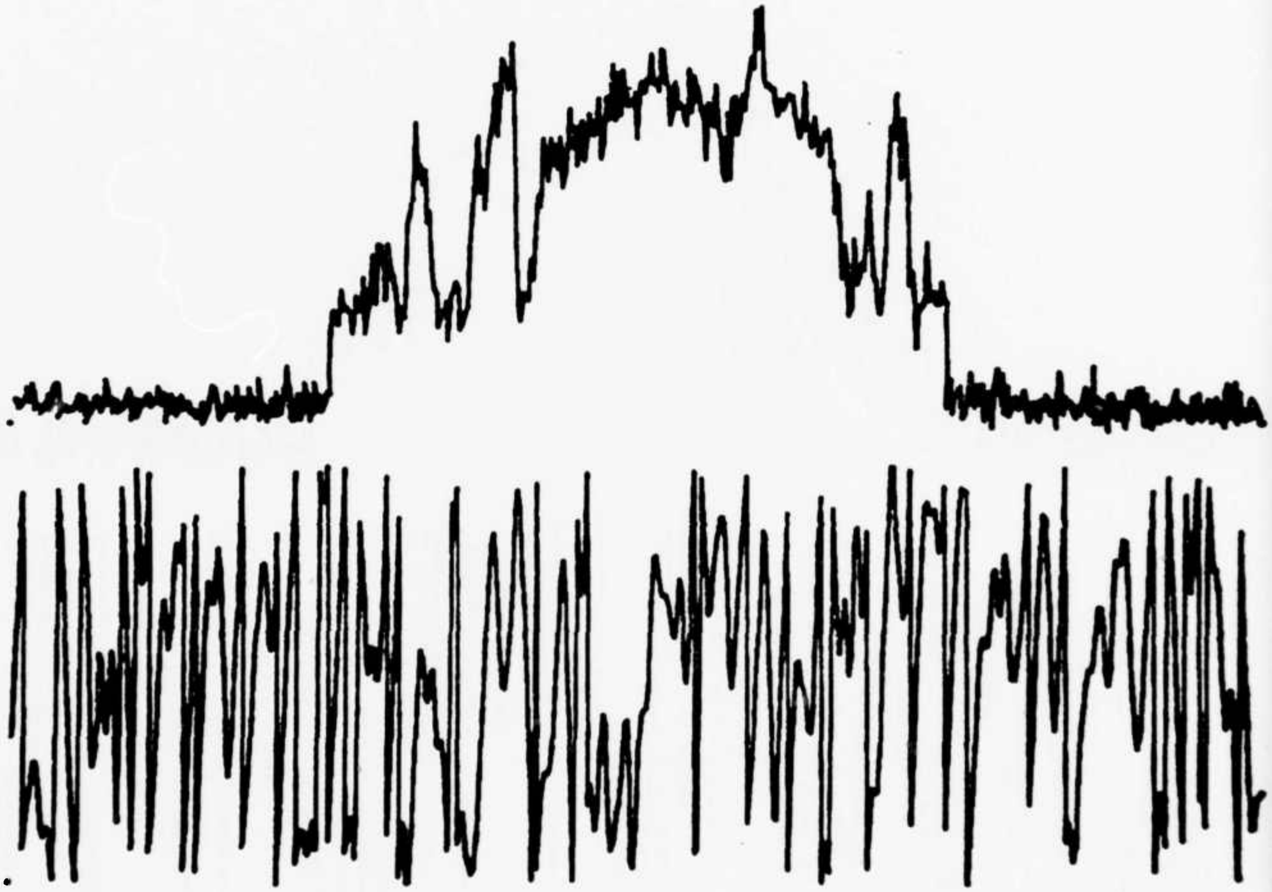


Figure 2. Image and Kinoform by using Hirsch's method

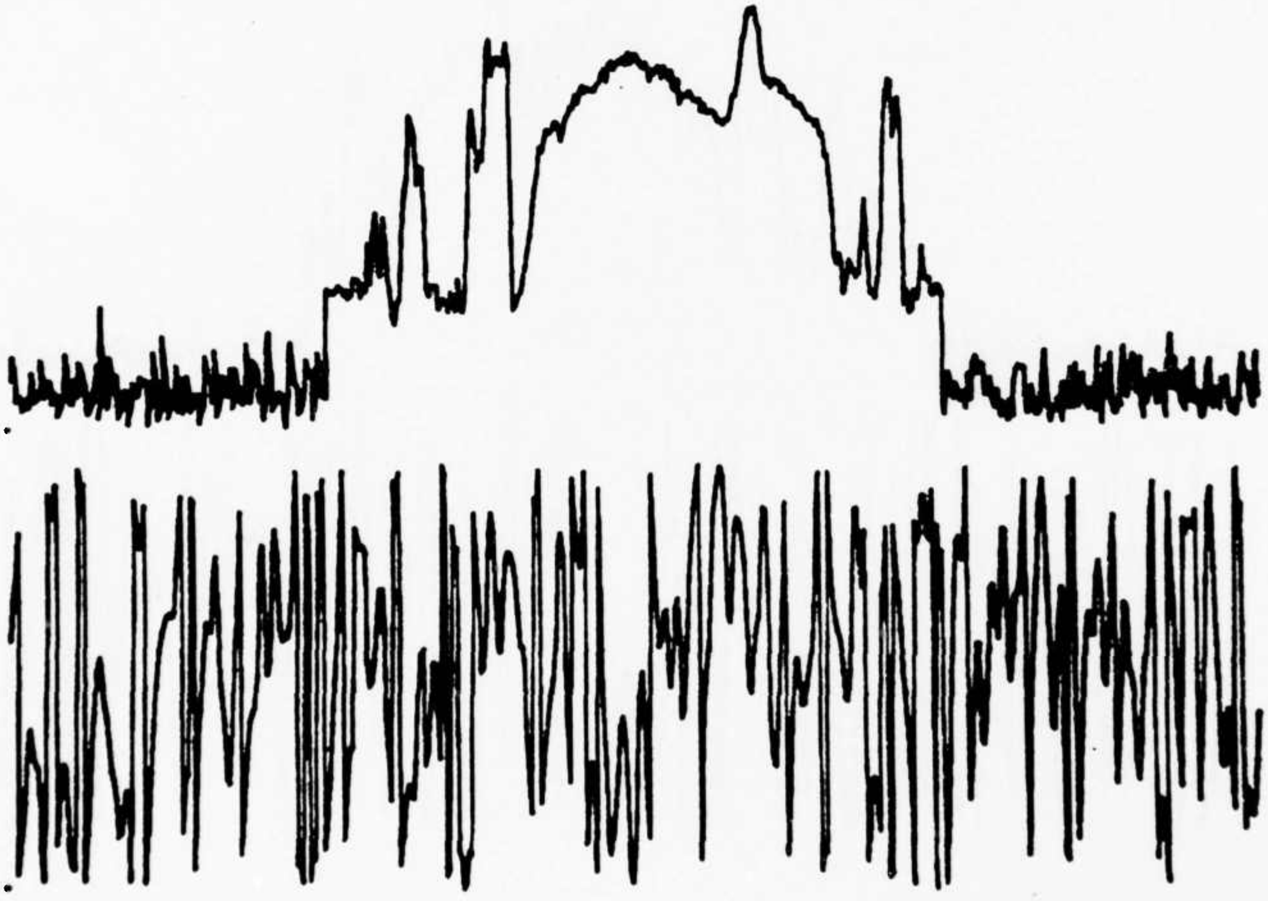


Figure 3. Image and Kinoform by using Fienup I

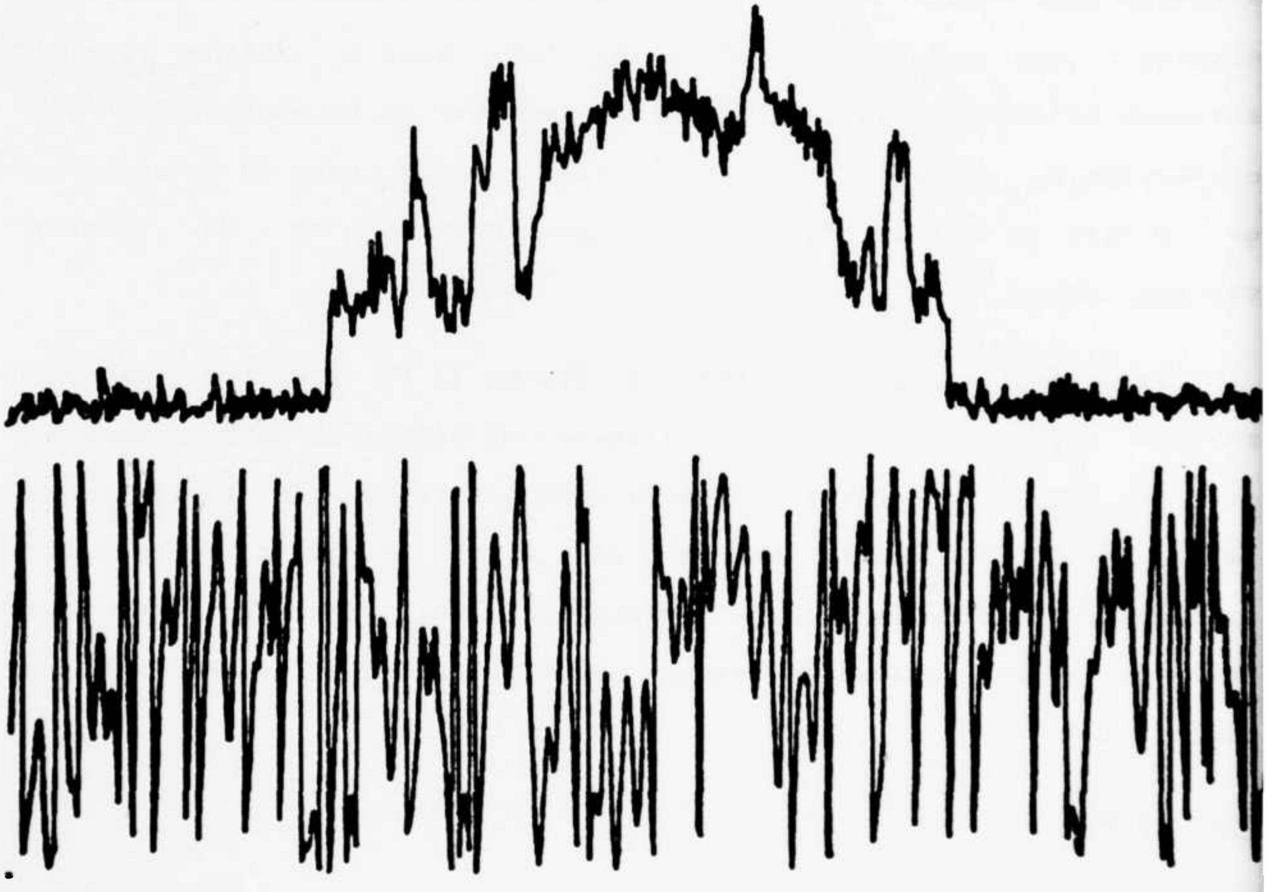


Figure 4. Image and Kinoform by using Fienup II

This data is consistent with the results shown in figures 2 to 4. Modification of the amplitude of the transform has also been tried. A biased cosine function is chosen because it is similar to the spectrum. The result is in figure 5 with error $E_4 = 0.07162$. Although this result is even worse than the first method, it does not eliminate the possibility of using this method. Further Fourier analysis is required to determine the component to be picked and its corresponding phase. Perhaps more than one component is required to get better results. This technique remains to be studied systematically.

Experiments so far indicate that Fienup II is the best of the methods tried. However, the performance of different methods changes from one object to another and since other methods are still under test, it is too early to conclude which method is the best for particular response functions. Future work will include tests and analysis of the various techniques and application to two-dimensional problems.

References

1. P.M. Hirsch, et.al., U.S. Patent No. 3,619,022, November 9, 1971.

2. N.C. Gallagher and B. Liu, "Method for Computing Kinoforms that Reduces Image Reconstruction Error," Applied Optics, 12, 1973, p. 2328.

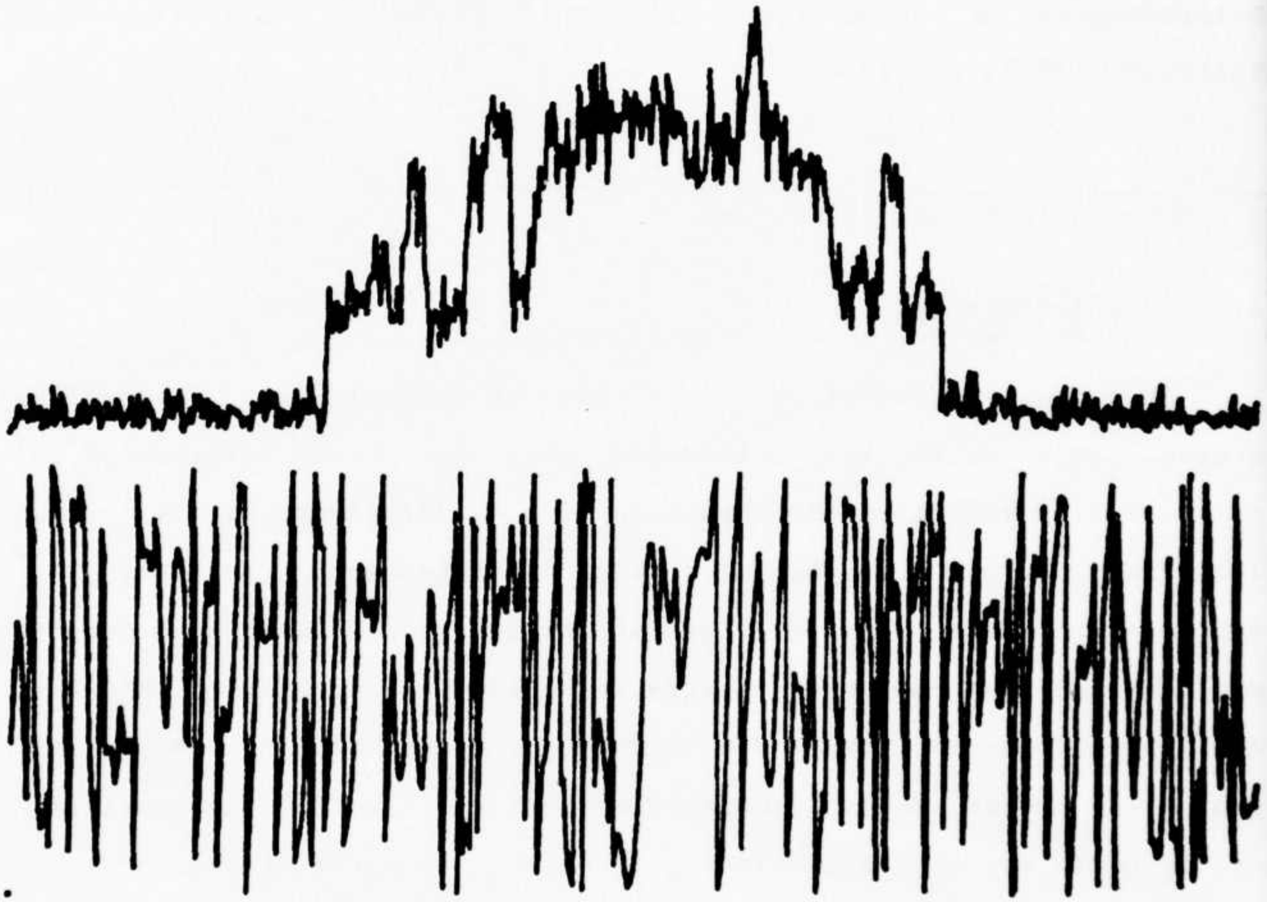


Figure 5. Image and Kinoform by using cosine function as the amplitude of the Kinoform

3. J.R. Fienup, "Improved Synthesis and Computational Methods for Computer-Generated Holograms," Ph.D. Thesis, Stanford University, 1975.

4. R.W. Gerchberg and W.O. Saxton, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," Optik, 35, 1972, p. 237.

3.5 ARTIFICIAL STEREO ALGORITHMS

Alexander A. Sawchuk

In a previous report, the technique of artificially generating stereo image pairs was introduced <1>. The stereo information is obtained from various image features or combinations of features, including pixel brightness, edge information, texture, or multispectral data. The technique is being explored as an enhancement aid for the human observer whereby these features are mapped into the depth perception ability of the observer. It should be emphasized that the stereo pairs generated are not reality, and the resulting stereo pairs may only coincidentally resemble the true stereo.

This section describes some of the details of the algorithm used to obtain the stereo and outlines current applications and extensions under exploration. Figure 1 is a block diagram of the entire process. To generate the artificial stereo pictures, feature extraction such as described above is performed on a single-frame monocular picture.

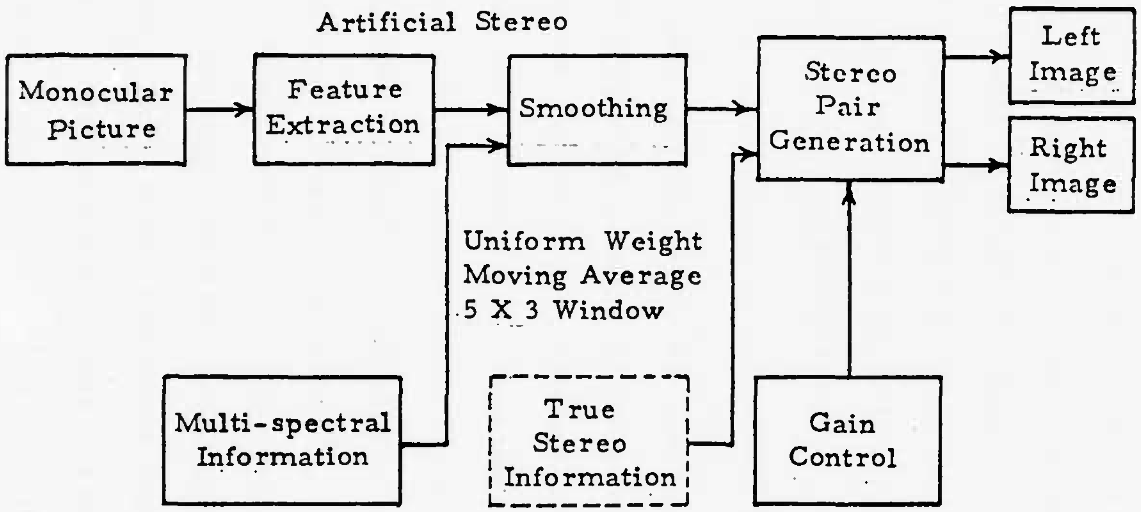


Figure 1. Artificial stereo block diagram.

Linear smoothing with a moving average convolutional window is then performed to eliminate any noisy or jagged edges on shifted parts of the stereo picture pair. The averaging window is necessary to avoid severe shape distortions which would make visual registration of local features difficult. The local features such as objects, spots and lines must remain for any stereo to be perceived. Various window sizes have been tried, and it appears that the optimum may be weakly picture-dependent. A good choice for most situations, however, is one with 3 pixel horizontal averaging and 5 pixel vertical averaging. The horizontal averaging is kept small so that blurring is minimized, and the vertical averaging is present to eliminate the most objectionable vertical jagged lines in shifted pixels.

The next stage of the operation is the actual stereo pair generation. Any true stereo information externally supplied can be introduced here, and a gain control for the stereo effect is provided in the algorithm. The stereo effect is obtained by local edge shifting as shown in figure 2. The perceived stereo in the left eye is obtained by locally shifting pixels to the right. For the right eye, the shift is in the opposite direction. Figure 3 shows the details of the shifting procedure for each line of the image. The top graph represents the desired non-negative height information after smoothing as a function of position along the image line. The second graph of figure 3 represents the input-output position characteristics. The height data is added to the unit slope linear function shown as a dotted line. With no height information, the dotted line would be a mapping with no shift from input to output.

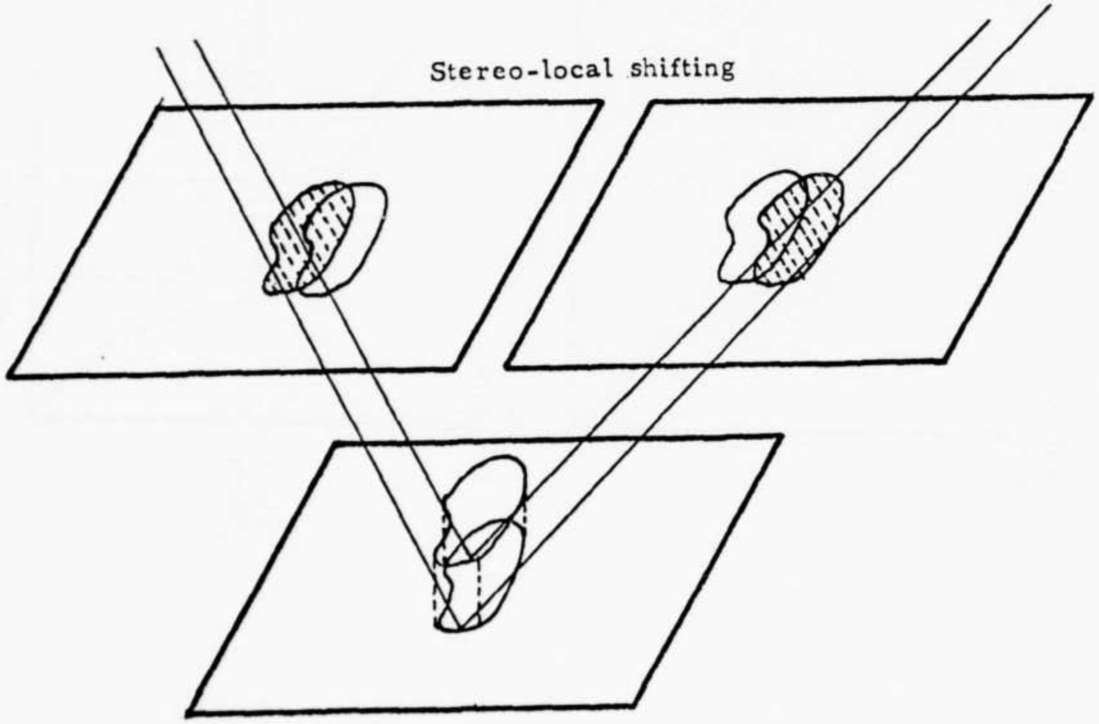


Figure 2. Local shifiting for depth illusion.

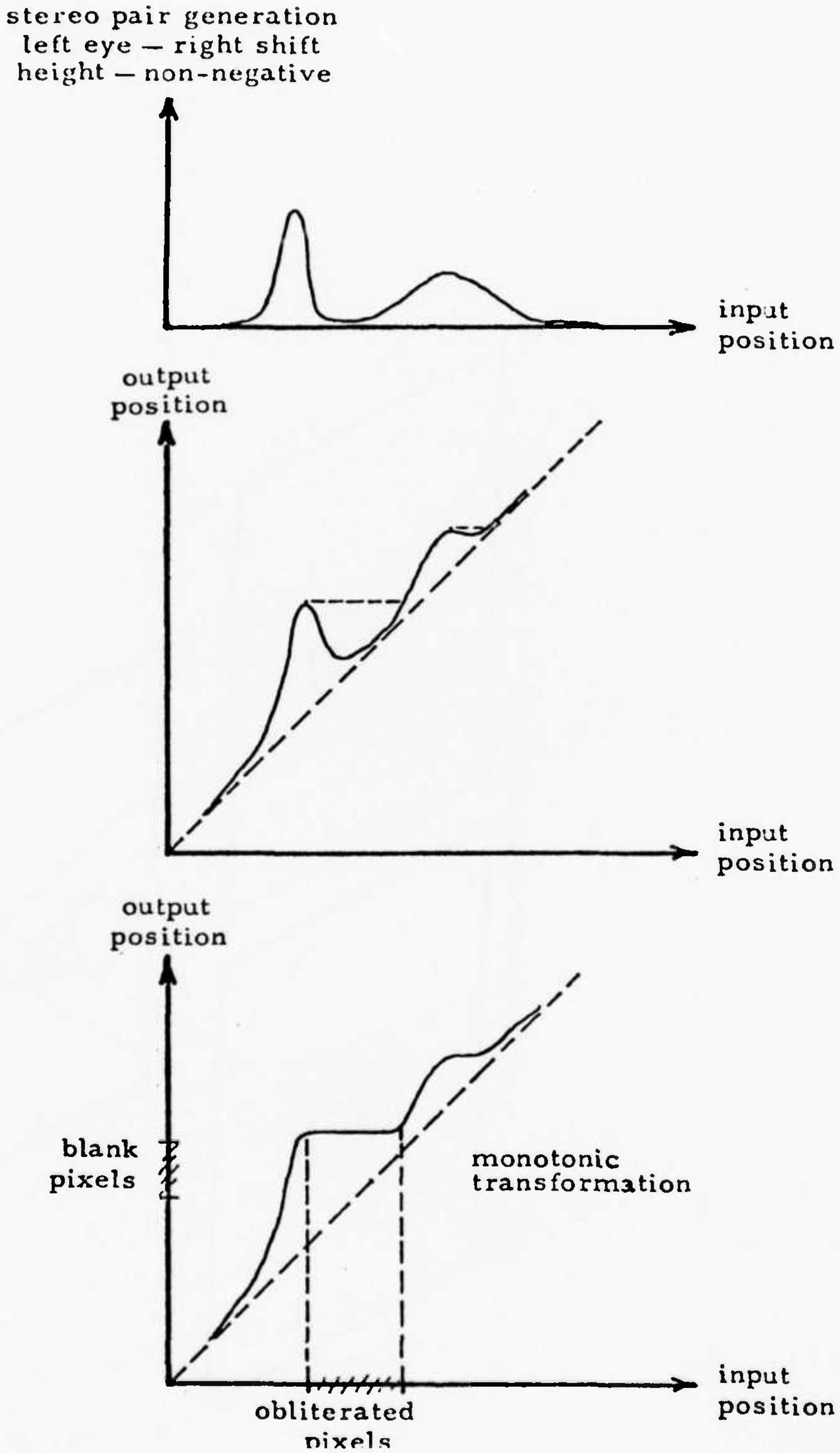


Figure 3. Shifting algorithm

The second graph is then made monotonic by a curve-following algorithm, as shown in the bottom graph of figure 3. In the input regions of zero slope, the pixels are obliterated by adjacent ones moved in. In regions of large positive slope, certain output areas may be missed and left blank. When this happens, pixels are moved in, beginning from left to right, to fill in the blank areas. Thus, a local distortion takes place to create the local shifting. The directions for the right eye are simply opposite.

Future work on artificial stereo will concentrate on the case of different features and algorithm improvement. It has been recently found that the addition of a brightness histogram equalizing step in the block diagram of figure 1 just before the stereo pair generation gives a much better distribution of heights and a more pleasing stereo illusion.

Reference

1. A.A. Sawchuk and H.C. Andrews, "Artificial Stereo from Image Features," USC Image Processing Institute Semiannual Technical Report, 1 September 1975 - 31 March 1976, USCIPI Report 660, March 1976, pp. 195-199.

4.0 SMART SENSOR PROJECTS

This section presents the results of the industrial liaison between USCIPI and Hughes Research Laboratories. The objectives of the past year's subcontract were twofold: first, to develop real time nonlinear optical data processing elements and second, to develop a CCD chip for smart sensor image understanding signal processing. Both objectives have been met, in part, with quite exciting results promised in the latter project. All projects are reported herein. The construction of the CCD chip is on schedule with completion, test, and evaluation expected within six months.

4.1 MOTIVATION FOR AN ADAPTIVE CCD IMAGE PROCESSING CHIP

Harry C. Andrews

This document represents some ideas in motivational thought which may be of some use in expansion of the Image Understanding CCD charge couple device signal processing chip work. The chip which was suggested for implementation under the first year subcontract had the Sobel operator amongst four other different operations as possible signal processing techniques to be considered for implementation. The motivation for these five different signal processing techniques may not be evident to the CCD implementors and therefore this report is designed to describe the thought processes that went into the generation of the mathematical formulation for the function suggested for chip implementation. The five operations so suggested are listed below.

1. The Sobel operator
2. The Mean operator
3. The Unsharp Masking operator which is equal to the Sobel operator plus a weighting of the Mean operator
4. The Thresholded Mean or Adaptive Binarizing operator
5. The Adaptive Stretching operator

These five operations are described mathematically in the attached statement of work.

The Sobel operator is essentially a convolutional kernel designed to detect edges. The operator requires a 3 x 3 convolving matrix which is designed to detect vertical edges and then designed to detect horizontal edges, the two outputs of the vertical and horizontal detectors then being combined by an absolute value operation to form the Sobel edge output. The absolute value operation is in lieu of a square rooting operation. This edge detector is a fairly useful one, requires only a small kernel; and in addition, contains nonlinear devices (i.e. the absolute value operations). The reason for selecting this edge operator is because of the nonlinearities involved as well as the possibilities for future implementation by looking at the ratio of the horizontal to the vertical edge energy. This will provide edge direction which also becomes a useful feature technique to be utilized in higher level image understanding operations. The Sobel operator can be viewed in image form and should provide a fairly clean edge detected image. References on this operation are readily available in the literature.

The mean operator is essentially nothing more than a low pass filter. A 3 x 3 matrix was suggested for implementation as a convolving kernel. The mean operator measures the average brightness in a 3 x 3 region surrounding the center pixel. This operator, as it stands alone, will tend to blur the image (remove high frequency noise as well as removing edge detail). The operator is suggested as possibly a signal processing low pass filter but more significantly as a switching signal to allow for adaptive signal processing in subsequent operations. Because the mean operator tends to track the low frequency variation in an image, it becomes very useful for discriminating between regions of high brightness and regions of large shadow or low brightness.

The unsharp masking operator is suggested as one technique to utilize with the mean operator. One could envision a weighting of a percentage of the Sobel output with a percentage of the mean output, the sum of the two percentages equaling 100%. Therefore, one could range from a completely edge image to a completely low pass or blurred image by modifying this particular percentage weighting factor. If this percentage weighting factor were in fact defined or determined by some information content within the image, then this would provide an adaptive signal processing aid. However, even if the percentage weighting factors are maintained constant or fixed throughout the image, this technique still qualifies for that under the general category of unsharp masking. Unsharp masking is a technique utilized by the Associated Press in transmitting many of their images across telephone communication lines. Unsharp masking essentially forms a low pass version of an image, subtracts it from the original image, weighted by a certain percentage such that the resulting image has higher frequencies emphasized by virtue of the fact that the low pass or blurred portion is removed.

The fourth signal processing function suggested for implementation is that known as the threshold mean or the adaptive binarizer. This function has numerous applications in a bandwidth compression mode of operation. Essentially the function is the following: the mean operator determines the local area mean and based on the local area, sets the center pixel or pixel in the center of the convolving kernel, to either completely on or completely off based upon whether that center pixel brightness is greater than or less than the mean average of its surrounds. This particular operation has the advantage of providing a binary video signal in the context of certain processes or has the advantage of providing binary adaptive thresholded signals as illustrated in adaptive binarization for fingerprint identification. The motivation in the latter application is in use of the a priori knowledge that the original fingerprint is best viewed under binary circumstances. However, latent fingerprints, when digitized, have a great many shades of gray and the adaptive binarizer has the property of tracking low frequency smudge contents and allowing edge and fingerprint information to be enhanced by subtracting off or essentially switching out the low frequency information. I might also suggest that another application for this particular operator is in the generation of binary video. As I understand it, in a lot of the Hughes Aircraft forward-looking infrared work, binary video becomes a very useful signal in two dimensions upon which to subsequently correlate. This particular type of binarization is adaptive in the sense that the local mean is detected by the mean operator and then controls binarization by tracking that mean in the bright region by rising high, and in low regions by lowering the amplitude values. The possible drawback of this particular binarizing filter lies in the fact that the filter is only 3 x 3 in extent. Possibly, the 5 x 5 and even 10 x 10 adaptive binarizing filters might be more useful. This would require a larger array on the chip but possibly because all of the weights are identical this is not too unfeasible a prospect.

The fifth and final application of the signal processing chip is that known as an adaptive stretching operator. Essentially this operation is also triggered based upon a decision by the local area mean detector. This operation is hypothesized by virtue of the fact that in regions that are fairly dark one would like to stretch the dark or low amplitude brightness values, and in regions that are fairly bright one would like to stretch the large or high amplitude brightness value. Therefore, this particular operation, known as adaptive stretching, is one in which the mean signal is tested against a constant half amplitude or half dynamic range brightness value. If in fact, the local area mean is greater than half amplitude brightness on the average, then the center pixel is stretched in the bright region. If in fact, the local area mean is less than half amplitude of the dynamic range possible, then the center pixel is stretched in the low brightness region. This operation will have the effect of adaptively selecting, based upon bright regicns or dark regions, and stretching the appropriate pixel value such that in dark regions, low amplitudes are enhanced and in bright regions, high amplitudes are enhanced.

In concluding this report and in considering the signal processing applications for the chip that has been proposed for development, one point I think is important to be made clear. Whether the actual signal processes that have been described become useful in an image understanding or higher level image processing environment, and whether the processed images themselves are pictorially and cosmetically more pleasing or not, is really not the point of the development of the chip under consideration. Rather, the selected five functions that have been suggested, themselves represent a large variety and great breadth of signal processing operations that conceivably could become quite useful on the front end of a sensor device. Therefore, it is possible that all the implementations will require larger kernel arrays or modifications of the signal processing operation described thus far. However, the important thing to emphasize is the fact that if these signal processing operations can indeed be implemented, and with the linear and nonlinear implications of the implementations so described, such implementation is a success even if the applications results are not striking and I feel that success in itself will motivate further investigation.

Appendix

I. There shall be two 3 x 3 convolving kernels on a chip.

kernel #1 shall be the Sobel operator, f

kernel 2 shall be the mean operator, f

II. There shall be various combinations of the outputs of the above kernels. These shall form five different output lines, and therefore output images.

1. Sobel operator = edge detection = $f_s(j,k)$
2. mean operator = low pass filter = $f_m(j,k)$
3. Sobel $+ \alpha$ mean = unsharp masking = $fu(j,k)$

4. Threshold mean = adaptive binarizer = $f_u(j,k)$

5. Function memory switch = adaptive stretching = $f_b(j,k)$

III. For every single image tested, there will be five images generated. The test procedure shall be either in real time TV or via digital eight bits 512 x 512 magnetic tape imagery.

Mathematically Defined Functions

1. $$f_s(j,k) = \frac{1}{8}\Big\{ \big| f(j-1,k-1) + 2\,f(j,k-1) + f(j+1,k-1) - f(j-1,k+1) - 2\,f(j,k+1) - f(j+1,k+1) \big| + \big| f(j+1,k-1) + 2\,f(j+1,k) + f(j+1,k+1) - f(j-1,k-1) - 2f(j-1,k) - f(j-1,k+1) \big| \Big\}$$

2. $$f_m(j,k) = \frac{1}{9}\Big\{ f(j-1,k-1) + f(j-1,k) + f(j-1,k+1) + f(j,k-1) + f(j,k) + f(j,k+1) + f(j+1,k-1) + f(j+1,k) + f(j+1,k+1) \Big\}$$

3. $$f_u(j,k) = (1-\alpha)\, f_s(j,k) + \alpha\, f_m(j,k) \qquad 0 \le \alpha \le 1$$

4. $$f_b(j,k) = \begin{cases} r & f_m(j,k) \le f(j,k) \\ 0 & f_m(j,k) > f(j,k) \end{cases}$$

$$5. \quad f_a(j,k) = \begin{cases} 2\ \mathrm{Min}\ \{f(j,k), \frac{r}{2}\} & f_m(j,k) \leq \frac{r}{2} \\ 2\ \mathrm{Max}\ \{f(j,k) - \frac{r}{2}), 0\} & f_m(j,k) > \frac{r}{2} \end{cases}$$

$f(j,k)$ = center pixel in 3 x 3 array to be processed

1 Sobel operator = edge detection

$$\begin{array}{cc} & j \longrightarrow \\ k \downarrow & \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \end{array} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$f_s(j,k) = \Big[\left| (a + 2b + c) - (g + 2h + i) \right| + \left| (c + 2f + i) - (a + 2d + g) \right| \Big] 1/8$$

output is center position given by $f_s(j,k)$

$$\begin{aligned} \text{i.e.} \quad f_s(j,k) = \Big[& \big| (f(j-1,k+1) + 2\,f(j,k-1) + f(j+1,k-1)) \\ & - (f(j-1,k+1) + 2\,f(j,k+1) + f(j+1,k+1) \big| \\ & + \big| (f(j+1,k-1) + 2\,f(j+1,k) + f(j+1,k+1)) \\ & - (f(j-1,k-1) + 2f(j-1,k) + f(j-1,k+1)) \big| \Big] 1/8 \end{aligned}$$

2 mean operator = low pass filter

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$f(j,k) = \frac{1}{9}(a+b+c+d+e+f+g+h+i)$$

output is center position given by $f_m(j,k)$

$$f_m(j,k) = \Big[f(j-1,k-1) + f(j-1,k) + f(j-1,k+1) + f(j,k-1) + f(j,k) + f(j,k+1) + f(j+1,k-1) + f(j+1,k) + f(j+1,k+1)\Big]\frac{1}{9}$$

3 Sobel + α mean = unsharp masking

$$f_u(j,k) = (1-\alpha)f_5(j,k) + \alpha f_m(j,k)$$

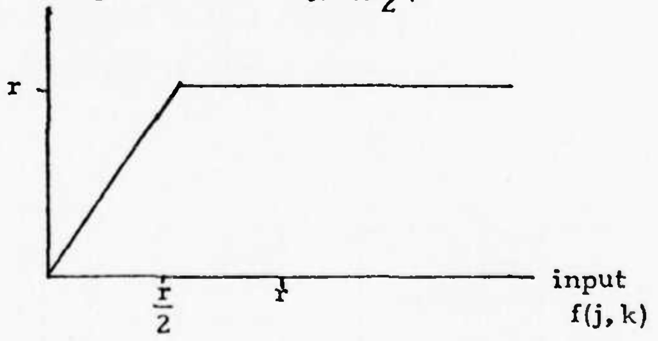
make α adjustable $0 \le \alpha \le 1$

4 Threshold mean = adaptive binarizer

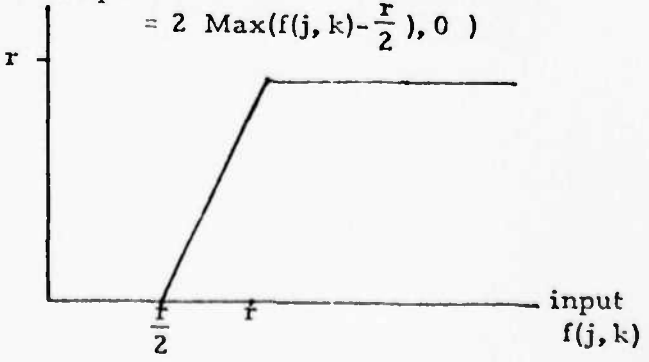
$$f_b(j,k) = \begin{cases} r & f_m(j,k) < f(j,k) \\ 0 & f_m(j,k) > f(j,k) \end{cases}$$

5 function memory switch = adaptive stretching

output $= 2\,\mathrm{Min}\,(f(j,k), \frac{r}{2})$



output $= 2\,\mathrm{Max}(f(j,k) - \frac{r}{2}), 0\ )$



r = input dynamic range

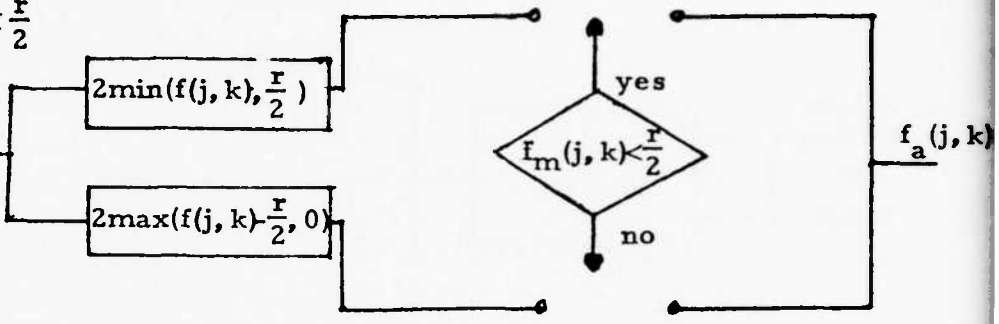
$\frac{r}{2}$ = half input dynamic range

The above transfer curves are to be switched between - depending on whether

$$\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} = f_m(j,k) \quad \text{is}$$

greater or less than r/2. The center pixel is to be the input

$$f_a(j,k) = \begin{cases} 2\min(f(j,k), \frac{r}{2}) & f_m(j,k) < \frac{r}{2} \\ 2\max(f(j,k) - \frac{r}{2}, 0) \end{cases}$$



4.2 CCD IMAGE PROCESSING CIRCUITRY

Graham Nudd

During the past year we have investigated the possibility of performing the following five algorithms with CCD/MOS integrated circuitry:

1. Chirp transformation
2. Robert's cross
3. Sobel operator
4. Hueckel operator
5. Histogram operation.

We have developed circuit concepts for each of the above (apart from the Hueckel operator) and include here block schematics together with a brief functional description. Details of the circuit design and layout are in general omitted, and only factors which directly affect the overall concept such as speed, dynamic range, etc. are discussed. We have also initiated work on the detailed design and layout of two

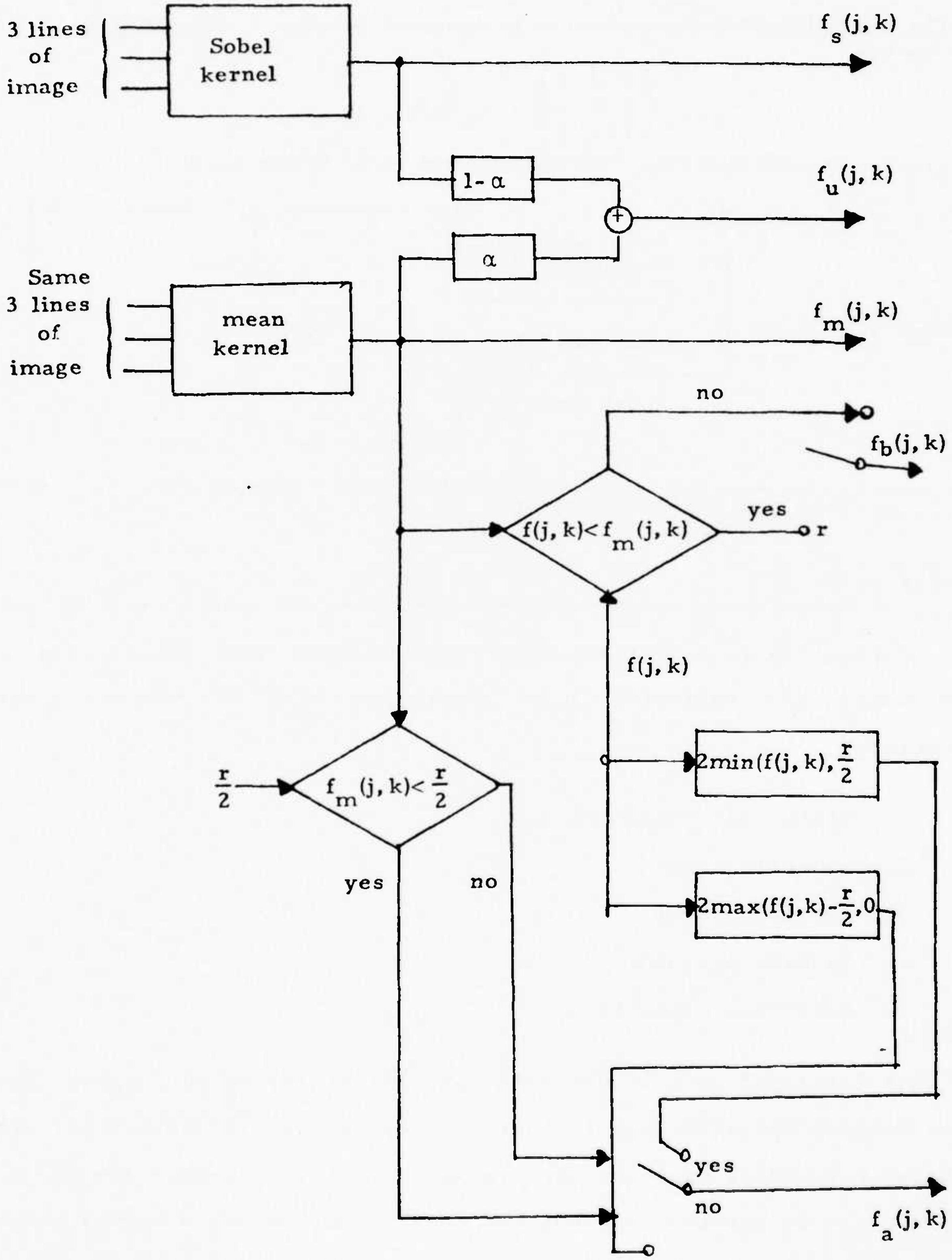


Figure 1. Total Configuration

test circuits, one of which (the Sobel operator) is currently being fabricated. Details of these are also included.

At the present time we have concentrated our efforts on charge transfer technology and compatible MOS circuitry, because of its inherently low power-delay product 10^{-2} pJ (versus 50 pJ for typical bipolar) and its application to both image detection and processing. This will be particularly advantageous for the development of "Smart Sensors." A concept of an integrated detector and pre-processor is shown in figure 1. Here a CCD imaging array is shown as detecting the image and creating charge carriers which are then clocked through a matrix of CCD/MOS processing circuits, (which might, for example, perform a two dimension convolution of the Sobel type) and the processed data stored in a CCD analog store.

Chirp Transformation

A one dimensional CCD implemented chirp transform has been described in the literature <1>. Factors affecting the accuracy and speed of this process are described in the previous semi-annual report. The major issues are in the area of (a) integrating the peripheral circuitry such as clocking circuits and input/output devices on the chip, (b) providing integrated accurate ($\geq$6 bit) multipliers. Both these areas are being pursued by other contractors, and hence we have devoted little effort in this area. The major impact Hughes Research Laboratories could have in this area is t employ its high resolution electron beam exposure techniques t fabricate a high resolution ($\geq$8 bits) CCD multiplying digital t

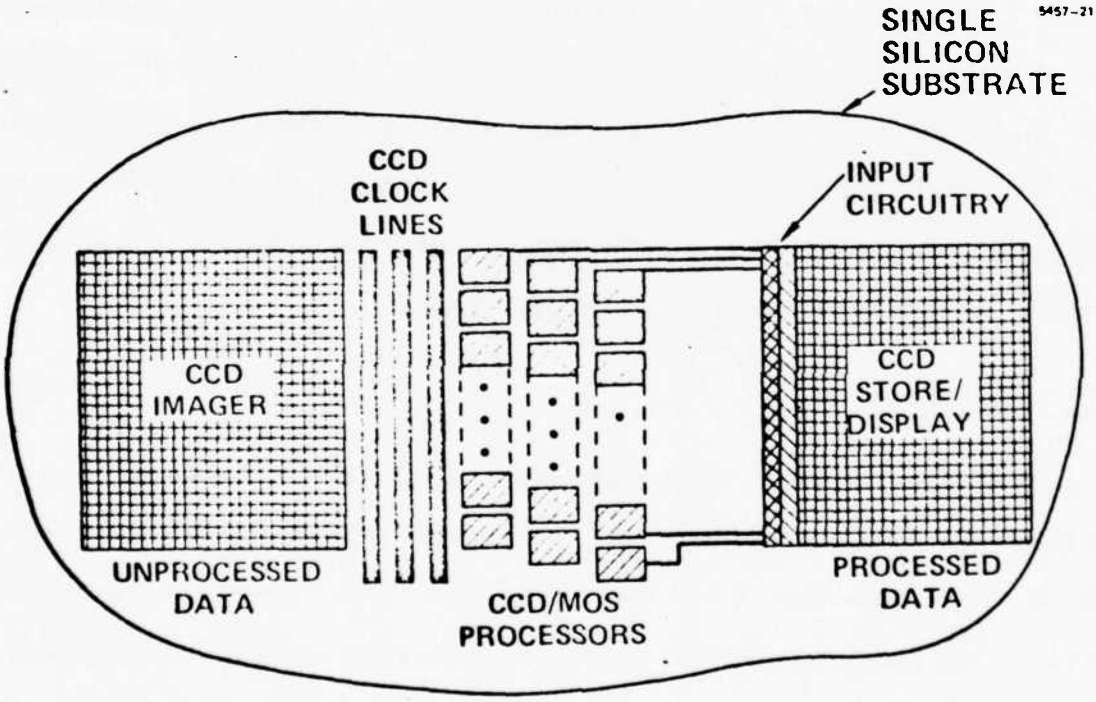


Figure 1. Concept of integrated CCD preprocessor.

analog converter <2> as previously discussed. Two dimensional Fourier Transform performed by concatenation of two one dimension transforms requires a permuting memory to transform the data. Work in this area is currently being sponsored by the Naval Undersea Center.

Robert's Cross Operator

A circuit which will perform the Robert's cross operator is shown schematically in figure 2. As shown it consists of two parallel CCD channels into which adjacent rows of picture elements are fed. The inputs (not shown in figure 2) might typically consist of the Tompsett surface potential equilibration circuits <3> or the CCD circuit might also be used as the sensor. The charge corresponding to the four adjacent pixels (a through d) used in the Robert's cross are simultaneously sensed by floating gate electrodes <4>. These outputs are used to drive a nonlinear CCD circuit which performs the magnitude operation $|a-d| + |b-c|$, required for the Robert's cross operation.

A schematic of the circuit is shown in figure 3. The two signals corresponding to pixels "a" and "d" are connected to parallel gates, as shown, and create potential wells beneath the oxide proportional to their magnitude. The source diffusions are then pulsed such that charge flows across the barrier ϕ_1 and into both wells. For surface channel devices the surface potential ϕ_s can be written as

$$\phi_s = V_o + V - \sqrt{V_o^2 + 2VV_o} \qquad (1)$$

where $V = V_g - V_{FB} - Q/Co$,

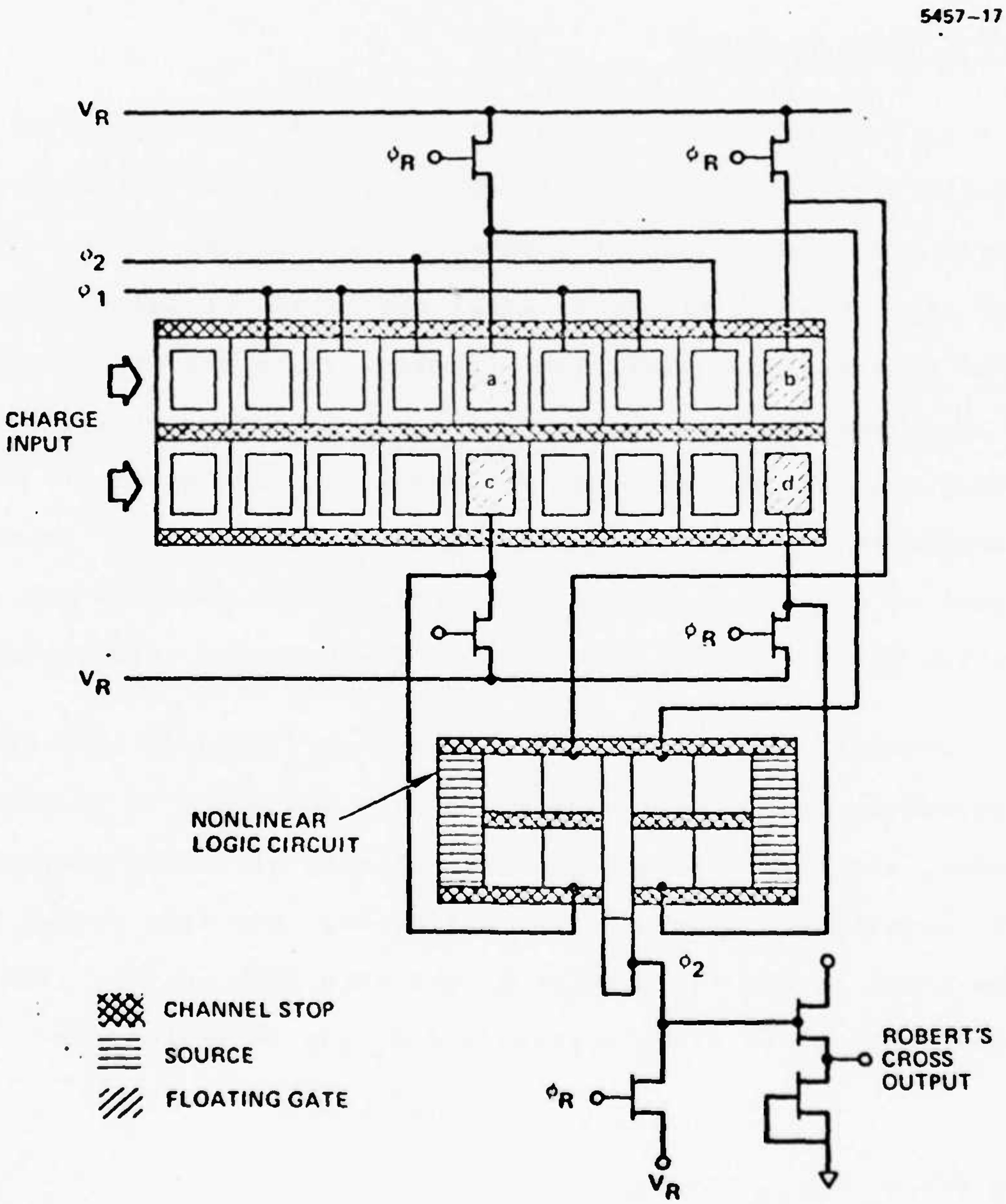
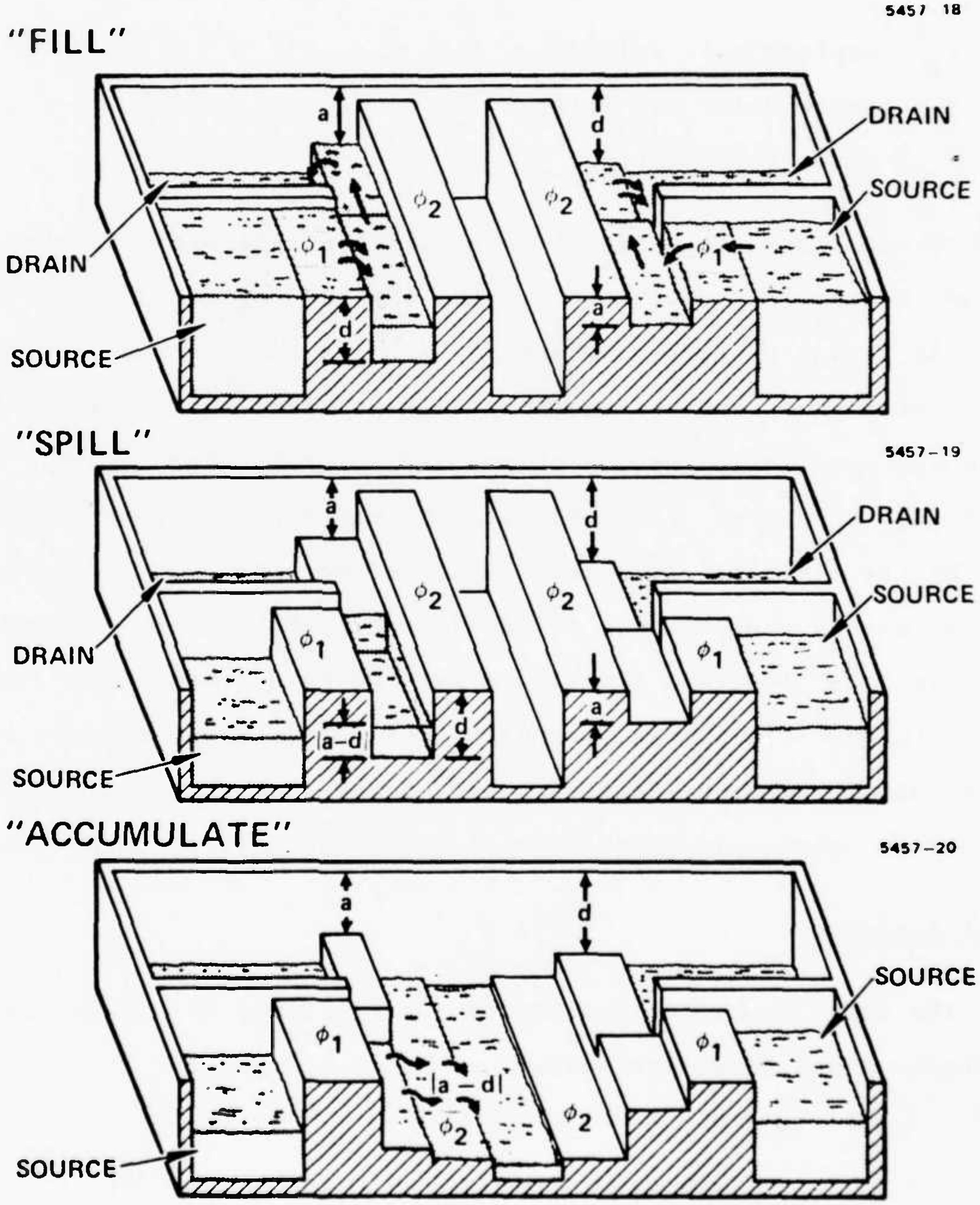


Figure 2. Circuit concept for Robert's cross operator.


"FILL"

5457 18

"SPILL"

5457-19

"ACCUMULATE"

5457-20

Figure 3. Schematic of nonlinear CCD circuit to evaluate |a-d|.

$$V_o = \frac{eq\,N_D}{C_o^2}$$

and

$V_g \equiv$ applied gate voltage

$Q \equiv$ charge under the gate

$C_o \equiv$ oxide capacitance

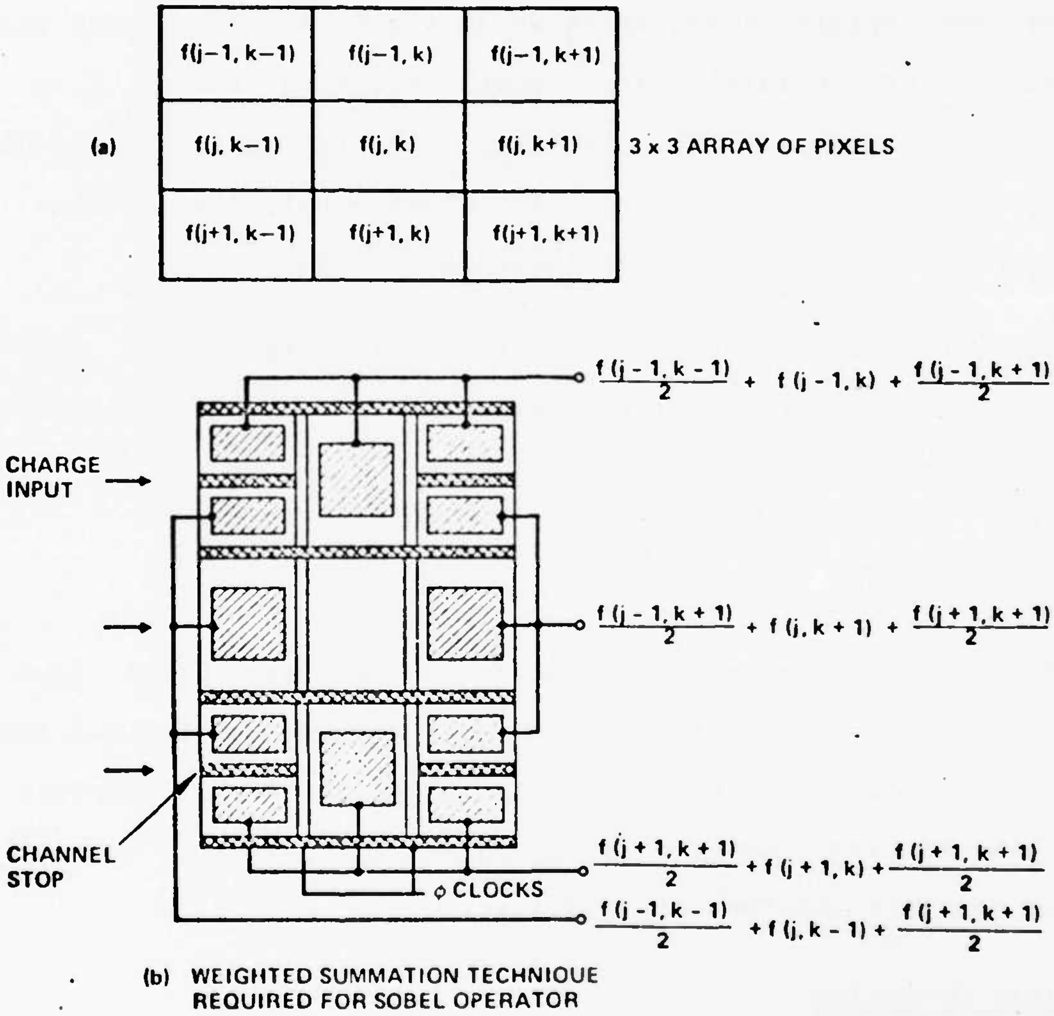
When charge flows the surface potential is uniform and hence the charge stored is proportional to the applied voltage ($Q = C_o V_g$). For the case shown in figure 3 where $|d| > |a|$ a charge proportional to (d-a) will be stored in the left half of the circuit. When the well shown has been filled excess charge will flow to the drain. Should $|a| > |d|$ a charge proportional to (a-d) would collect in the right side of the circuit. In either case after the source voltage is lowered charge proportional to $|a-d|$ is dumped into the central bucket as shown. An identical circuit is used to form $|b-c|$ and the full output $|a-d| + |b-c|$ can then be sensed. The output circuit is shown as a simple source follower and reset but techniques such as correlated double sampling can be used, if necessary, to reduce noise.

Sobel Operator

The Sobel operator operates on a 3 x 3 array of pixels as shown in figure 4 and forms the output

$$\begin{aligned} f_s(jk) = \frac{1}{8} \Big\{ &\big| f(j-1, k+1) + 2f(j, k-1) + f(j+1, k-1) \\ &-f(j-1, k+1) + 2f(j, k+1) + f(j+1, k+1) \big| \\ &+ \big| f(j+1, k-1) + 2f(j+1, k) + f(j+1, k+1) \\ &- f(j-1, k-1) + 2f(j-1, k) + f(j-1, k+1) \big| \Big\} \end{aligned}$$

5457–15



(a)

f(j–1, k–1)	f(j–1, k)	f(j–1, k+1)
f(j, k–1)	f(j, k)	f(j, k+1)
f(j+1, k–1)	f(j+1, k)	f(j+1, k+1)

3 x 3 ARRAY OF PIXELS

CHARGE INPUT
CHANNEL STOP
ϕ CLOCKS
$\frac{f(j-1,k-1)}{2} + f(j-1,k) + \frac{f(j-1,k+1)}{2}$
$\frac{f(j-1,k+1)}{2} + f(j,k+1) + \frac{f(j+1,k+1)}{2}$
$\frac{f(j+1,k+1)}{2} + f(j+1,k) + \frac{f(j+1,k+1)}{2}$
$\frac{f(j-1,k-1)}{2} + f(j,k-1) + \frac{f(j+1,k+1)}{2}$

(b) WEIGHTED SUMMATION TECHNIQUE REQUIRED FOR SOBEL OPERATOR

Figure 4. Weighting technique for Sobel Operator.

Two basic operations are involved (a) nondestructive charge tapping and weighting and (b) magnitude operation and summation similar to that described for the Robert's cross:

(a) The weighted summation can be performed by varying the area of the sensing gates as shown in figure 4. Here three parallel analog CCD registers are shown receiving charge from three adjacent rows of the picture, and the gates sense an output proportional to their area. For clarity cnly the storage gates are shown and the clocking lines are omitted.

(b) The magnitude operation can be achieved by a number of techniques similar to that described for the Robert's cross and the summation by charge merging.

Hueckel Operation

During the course of this study we have spent some time investigating the Hueckel operator. Our general conclusions are that analog CCD processing cannot directly be exploited to perform this algorithm in its present form, and we have therefore not developed circuit concepts for this at this time.

Histogram Operation

A block schematic of the circuit elements required to perform a real-time histogram are shown in figure 5. In principle, the input data is used to set up a potential barrier $\emptyset_b$ at the surface of the CCD and a linear current source is used to accumulate charge adjacent

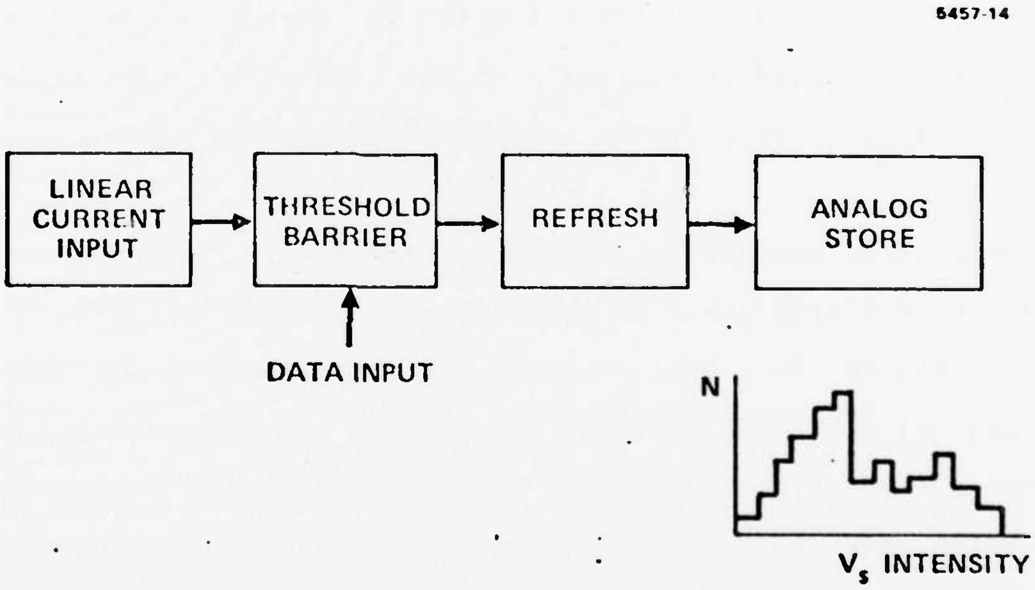


Figure 5. Block schematic of histogram circuit.

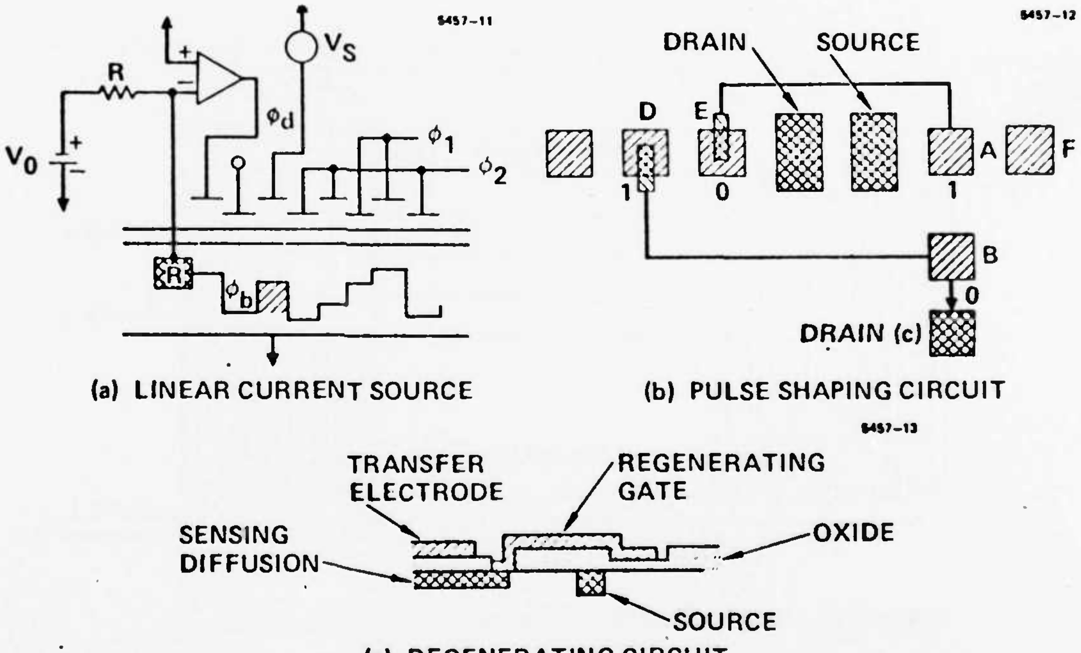
to this barrier. When the charge stored is sufficient to overcome current flows thereby converting amplitude to time delay. The charge crossing the $\emptyset_b$ triggers a pulse shaping network which feeds data into an analog store thus forming the histogram. The circuit which provides the linear current source is shown in figure 6a. Here the first stage of the CCD acts as a virtual drain and the current is self-regulating via the feedback path around the amplifier. The charge stored after time t is

$$Q(t) = \frac{V_o}{R} t \quad .$$

For an input signal V_s causing a barrier $\emptyset_b$ charge flows when $Q(t)$ $C_o \times V_s$ or at time $t = (V_s/V_o)(C_oR)$. If the shift register is clocked at f_c an output signal appears after n clock cycles where

$$n = \frac{f_c C_o R}{V_o} V_s \quad .$$

This data, after some pulse shaping, is fed into a serially accessed digital CCD register prior to parallel readout into an analog store as shown in figure 7. The accuracy of the circuit is determined primarily by the noise in the input circuit and error accumulating in the analog store. The equivalent noise charge in the input circuit is $\sqrt{KTC}$ which for 15 x 50 $(\mu m)^2$ gate size will be approximately 150 electrons. For 8 bit quantization the charge storage must therefore be in excess of 80 x 10^3 electrons, which implies a gate voltage swing of >1 V. Since the current from the input circuit shown in figure 6a will be of the order of 10^{-9} A the pixel data rate is limited to $f_s \leq 100$ kHz. requiring the amplifier to settle in approximately 4 x 10^{-8} seconds. Discrete devices are currently available at these rates in which case the digital CCD register would be required to



(a) LINEAR CURRENT SOURCE

(b) PULSE SHAPING CIRCUIT

(c) REGENERATING CIRCUIT

Figure 6. Circuit elements required for histogram circuit.

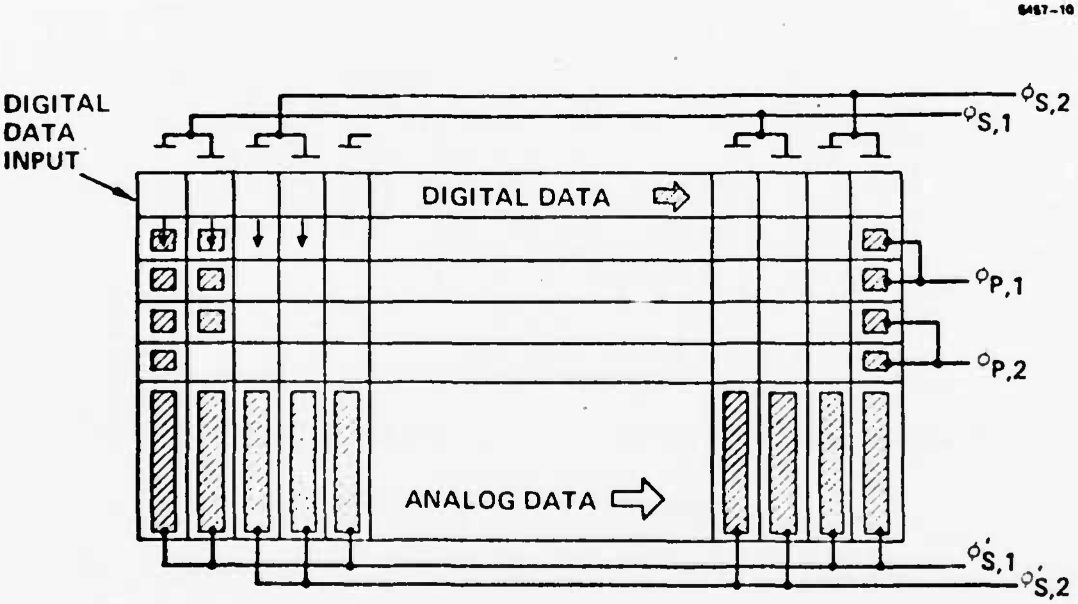


Figure 7. Digital CCD register and analog store.

operate at approximately 25 MHz, which is about the present state of the art. Data from this register is shifted out in parallel to the analog store (figure 7) which accepts data at the pixels rate, 100 kHz. The total storage time is practically limited to about 0.1 seconds by dark current, limiting the operation to a total of about 10^4 samples.

The pulse-shaping circuit illustrated in figure 6b and 6c must be included to provide a logical one to the input of the digital shift register. As shown, the circuit is based on digital refresh described by Kosnocky and Carnes <5>. After charge flows over the barrier $\emptyset_b$ a string of logical ones with charge equivalent to approximately 10^5 electrons will flow. The circuit uses floating gate sensors to provide a single pulse output coincident with the start of this string. When data under gates D and E are both "zeros" the refresh circuit will create potential wells under both gates A and B. This will cause data to flow to drain C. When the input data changes (from zero to one) the potential well under gate B disappears and the charge drawn from the source is clocked through to F prior to entering the CCD digital delay. When both D and E are "ones" the potential under gate A rises and no charge flows thus providing a single data output at a time equivalent to the start of current flow. A cross section of one of the gate interconnects is shown in figure 6c together with the time constants involved. Present photolithographic techniques indicate that this circuit will operate with delays of less than a few nanoseconds.

Test Circuits

To emphasize that the signal processing operations described here can indeed be implemented on single small chips using CCD/MOS circuitry, we are presently supporting, with cur own internal research funds, the development of test circuits.

Sobel Operator

As shown in figure 8, the CCD Sobel operator consists of a two dimensional CCD filter, CCD absolute value circuits and a charge summer. The input format is three parallel or demultiplexed channels of neighboring scan lines on the image plane. The sampled analog charges of a 3 x 3 detector submatrix are nondestructively transferred, sensed, added, and subtracted in a mask-programmed CCD. The outputs of the two dimensional CCD filter are then fed into the CCD absolute value circuits where charges corresponding to the input magnitudes are generated. The final output is the charge sum of the absolute values and can be converted into voltage forms for further processing.

The above described CCD Sobel operator has been designed on a single small chip which might easily be integrated with a focal plane detector array, thus forming an integrated sensor-processor chip which might have an exciting and important impact on future image processing systems. This circuit is currently being fabricated and we anticipate operating devices in several months.

Test Circuit II

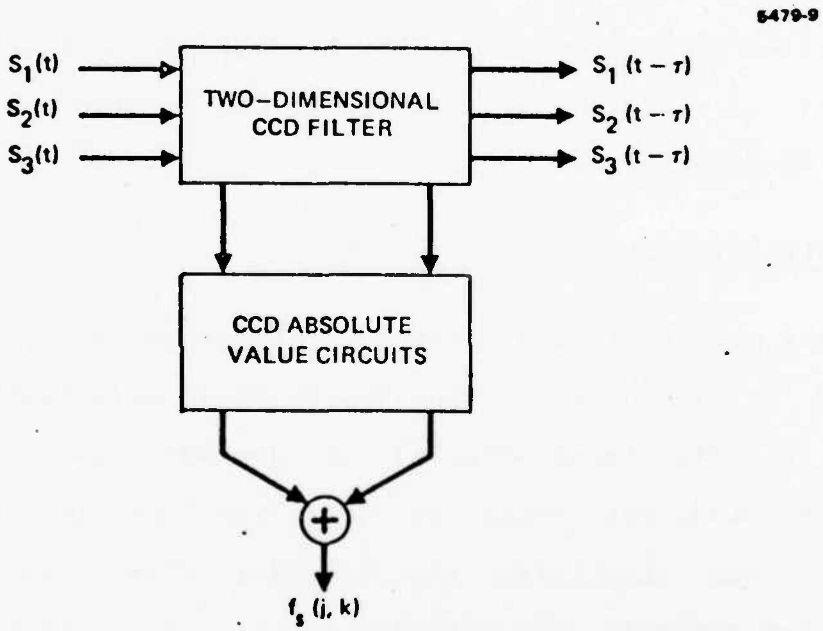


Figure 8. Schematic of Sobel Operator.

Five basic algorithms are addressed in the second test circuit:

1. Low pass filtering
2. Sobel edge detection
3. Unsharp masking
4. Adaptive stretching
5. Binarizing

A block schematic of the algorithm is shown in figure 9a together with the five separate outputs. It is designed to operate on a subarray of 3 x 3 adjacent picture elements, and a simplified circuit concept is illustrated in figure 9b. A brief description of each circuit component is given below.

Low Pass Filtering

The low pass filtering operation is achieved by the parallel connection of nine adjacent floating gates as illustrated at the left of figure 10. Here three parallel CCD channels are shown, through which three adjacent lines of data are fed (as for the Sobel operator). For simplicity the transfer gates and clocking and circuitry are omitted. The weighted outputs are connected to a single output line which drives a source follower shown. The load is a variable conductance consisting of a FET with variable gate voltage providing an output

$$\begin{aligned} f_m(j,k) = \frac{\alpha}{9} \Big| & f(j-1,\ k-1) + f(j-1,k) + f(j-1,\ k+1) \\ & + f(j,k-1) + f(j,k) + f(j,\ k+1) \\ & + f(j+1,\ k-1) + f(j+1,k) + f(j+1,\ k+1) \Big| . \end{aligned}$$

The tapping technique is nondestructive and thus, after sensing, the

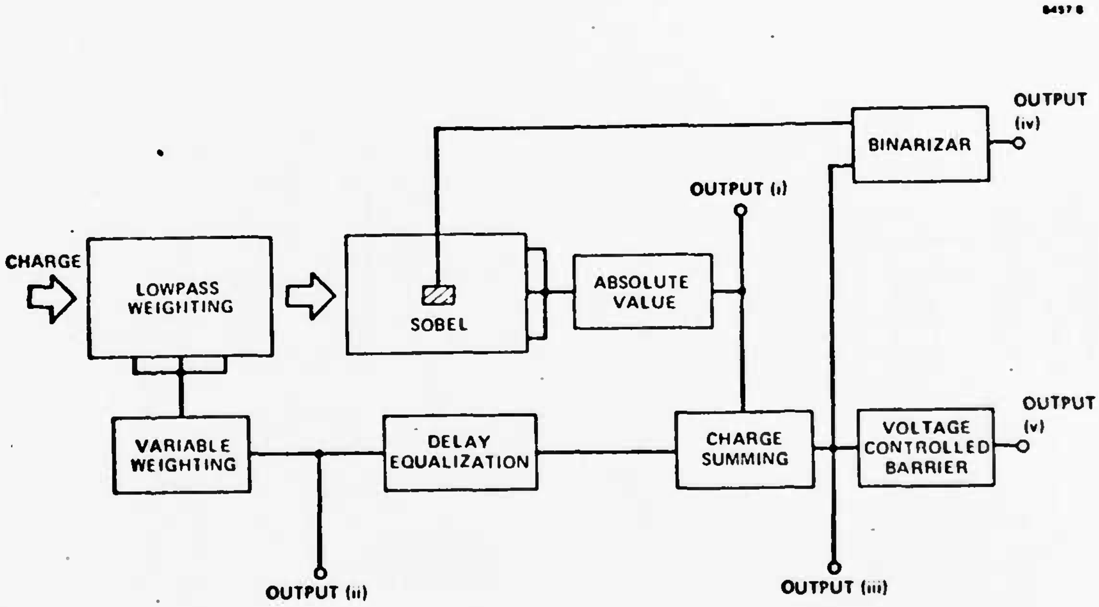


Figure 9(a). Block schematic for test Circuit II.

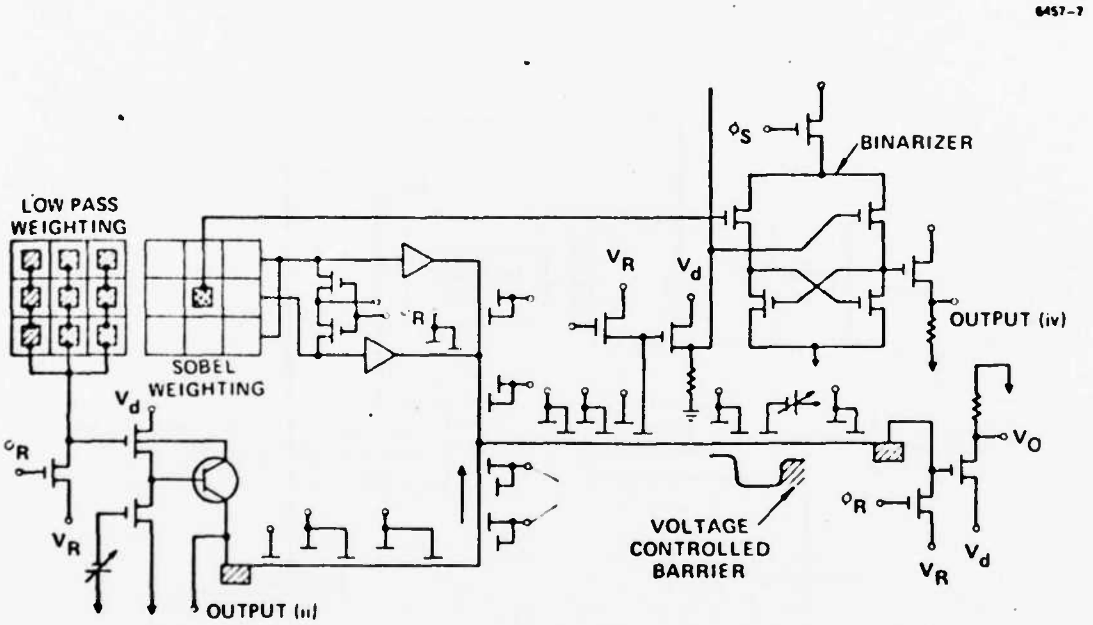


Figure 9(b). Circuit concept for test Circuit II.

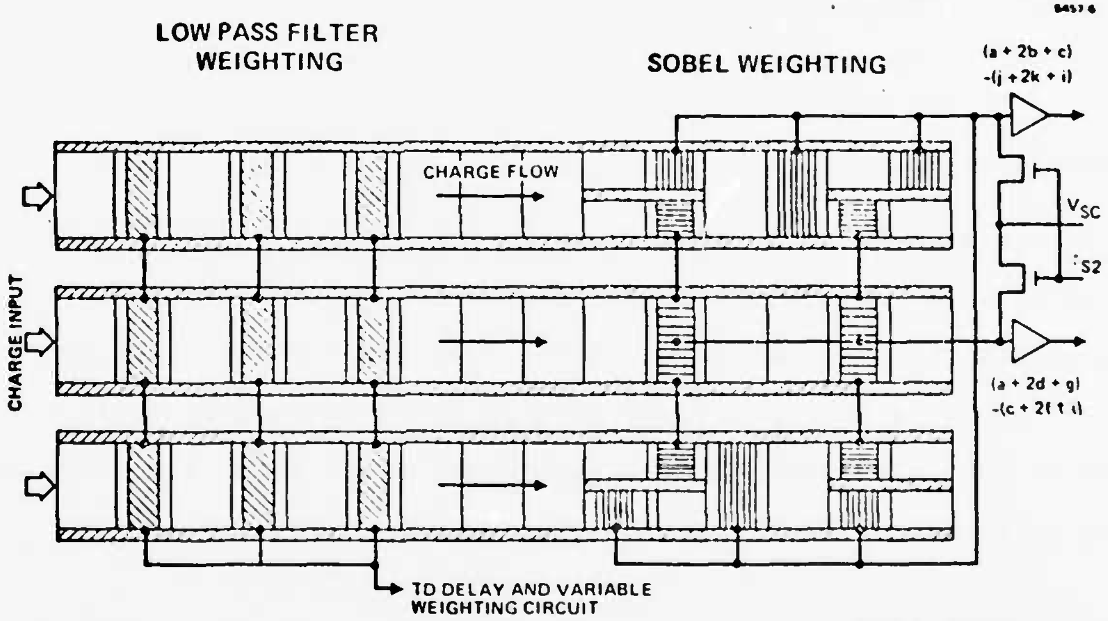


Figure 10. Weighting required for test Circuit II.

charge is clocked through to the Sobel operator.

Sobel Operator

The weighting required for the Sobel operator $f_s(j,k)$ is shown on the right of figure 10 (again the transfer gates and clocking circuitry are omitted for clarity). It is essentially identical to that described under Test circuit I. However, the input charge will be clocked directly from the low pass filter as shown and the output will be fed to the charge summing circuitry shown in figure 11.

Unsharp Masking

The unsharp masking output $f_u(j,k) = f_s(j,k) + \alpha f_m(j,k)$ is formed by the charge summing circuit shown in figure 11. Here charge corresponding to the low pass filter and the Sobel operator are clocked under a single storage gate as shown tc form $f_u(j,k)$. Three stages of CCD delay are shown in the output path from the low pass filter to equalize the charge delay prior to summation.

Adaptive Stretching

The unsharp masking output is further clocked through the CCD to a barrier ($\emptyset_b$) which is controlled by an externally applied voltage. If the signal charge Q_{usm} is less the $\emptyset_b/C_{ox}$ the charge is collected on the output drain. Charge equivalent to $Q_{usm} - \emptyset_b/C_{ox}$ is clocked over the barrier to the remaining circuit. The effect of this is to apply a threshold prior to a variable gain output connected to the adaptive stretching pad shown.

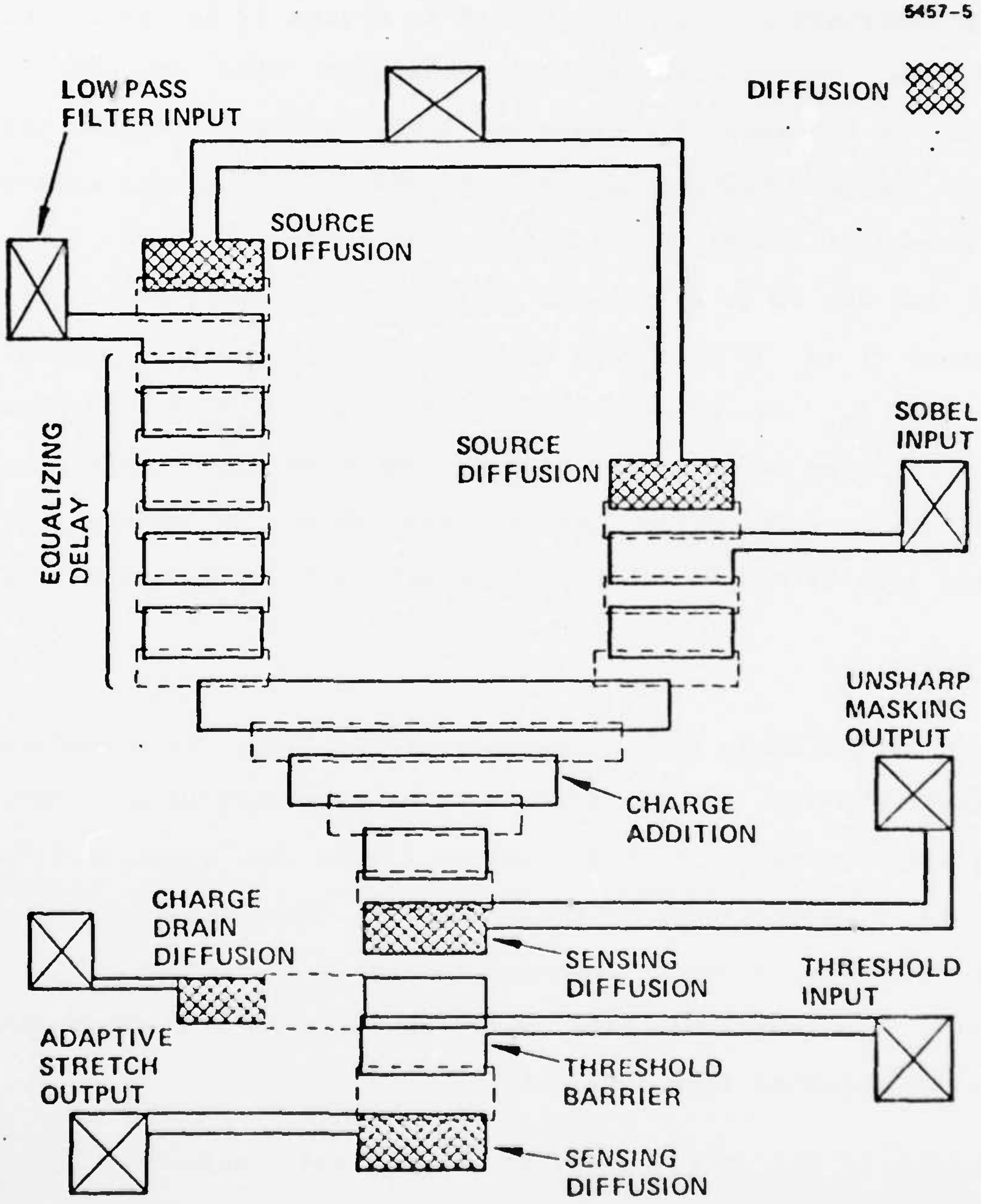


Figure 11. Illustration of layout for unsharp masking and adaptive stretching.

Adaptive Binarizer

The binarized circuit illustrated in figure 12 is based on the regenerative charge comparator techniques used in CCD digital memories. It consists of a cross coupled flip flop. Charge entering from the Unsharp Masking operator and the center picture element are sensed at output 01 and 02, respectively. The voltage difference between the two nodes will cause differences in the conductivities of transistors T1 and T2 which act as loads for the cross coupled flip flop. When $\emptyset_d$ is pulsed the gate voltages on T3 and T4 therefore differ, resulting in a further current imbalance which rapidly shuts one transistor off while driving the other to saturation. The resulting zero or one is taken out through load devices (not shown).

Future Plans

Work on the above test circuits will continue. It is anticipated that test devices for Circuit I will be available by year end, when testing will commence. We have devised a test plan which will use an IMSAI 8080 microprocessor to access U.S.C. data tapes and also drive the CCD circuits. A block schematic is shown in figure 13. After processing the supplied data will be displayed as shown and also supplied on magnetic tape to U.S.C.

Because of the wide variety of linear and nonlinear operations involved in Circuit II we consider that a demonstration of the feasibility of this circuit is most important and we have large impact in the community. We therefore propose that this circuit be developed

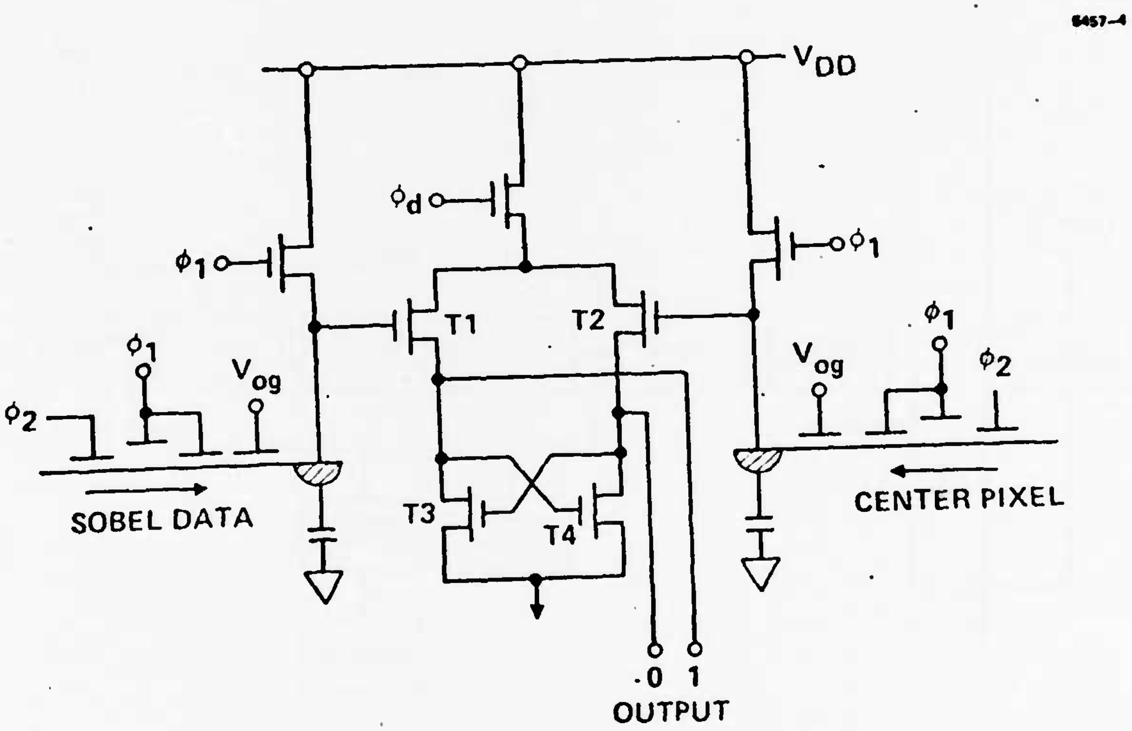


Figure 12. Concept of binarizer.

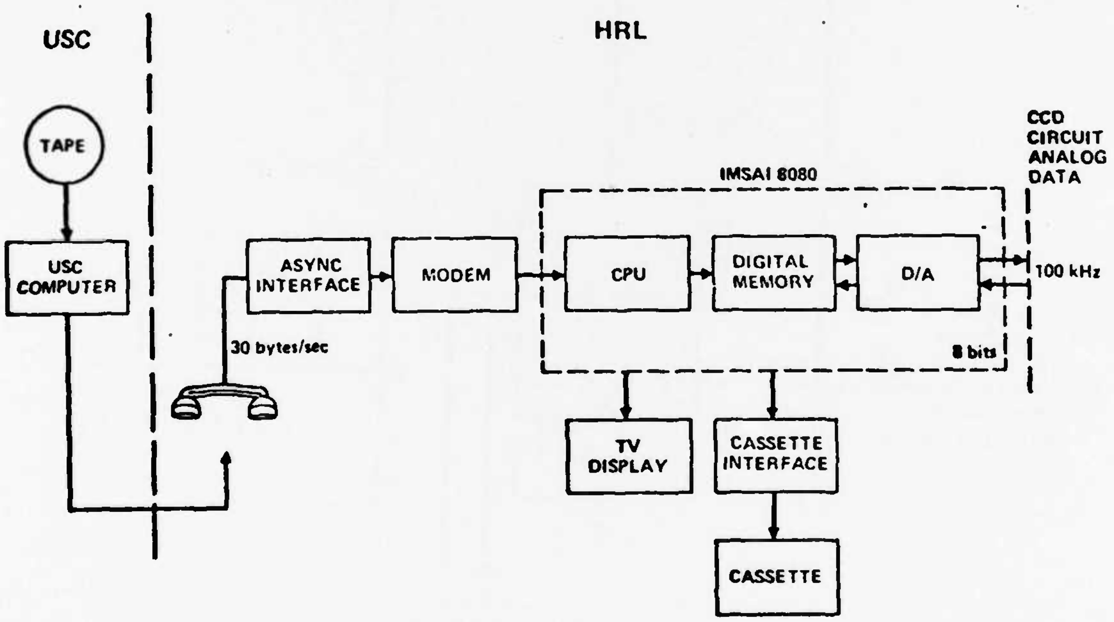


Figure 13. Schematic of test setup.

as part of the next phase of this program.

References

1. R.W. Brodersen, C.R. Hewes, and D.D. Buss, "A 500-Stage CCD Transversal Filter for Spectral Analysis," IEEE ED-23, No. 2, February 1976, p. 143.

2. J.P. Albaran and D.A. Hodges, "A Charge Transfer Multiplying D to A Converter," 1976 IEEE International Sclid State Conference, Philadelphia, February 1976.

3. M.F. Tompsett, "Surface Potential Equilibrium Method for Setting Charge in CCDs," IEEE ED-22, No. 6, June 1976.

4. M.H. White, I.A. Mack, F.J. Kirk, D.R. Lampe, and J.L. Fagan, "An Analog CCD Transversal Filter with Floating Clock Electrode Sensor and Variable Tap," IEEE First Solid State Conference, Philadelphia, February 1976.

5. W.F. Kosnocky and J.E. Carnes, "Charge-Coupled Digital Circuits," IEEE, Vol. SC-6, No. 5, October 1971.

4.3 REAL TIME NONLINEAR OPTICAL DATA PROCESSING

J.L. Erickson, A. Au and J. Grinberg

During the past year we have been using the hybrid field effect liquid crystal light valve (LCLV) as an incoherent-to-coherent interface in optical data processing (ODP) experiments.

Simultaneously, the LCLV has functioned as the thresholding/saturating (clipper) device required in order to achieve nonlinear ODP functions as reported by S.R. Dashiell and A.A. Sawchuk (Optical Communications, 15, 66-69, 1975).

The structure and operation of the hybrid field effect LCLV are described in detail in "A New Real-Time Noncoherent to Coherent Light Image Converter - the Hybrid Field Effect Liquid Crystal Light Valve," J. Grinberg, et. al., Optical Engineering, 14:3, 217-225, May/June, 1975. This device was developed for linear OPD applications, and so is not optimized for the nonlinear ODP process.

Our approach has been to study the characteristics of the LCLV in a level slicing experiment. Level slicing is a difficult nonlinear ODP function to implement because it requires a very high gamma response, rapidly saturating clipper. This is a useful approach, because the results can be easily interpreted in terms of the device characteristics.

The experimental setup uses the coherent optical data processing system shown in figure 1 with variable definitions. The projection light is derived from a mercury arc lamp and filtered at a center wavelength of 5250 angstroms. The coherent light source is a 50 mW He-Ne laser. The densitometry curve for the light valve operating at 4.0 volts RMS at 2 KHz was measured and plotted in figure 2. Maximum gamma was 2.1.

Figure 3 is the transfer function achieved in a level slicing

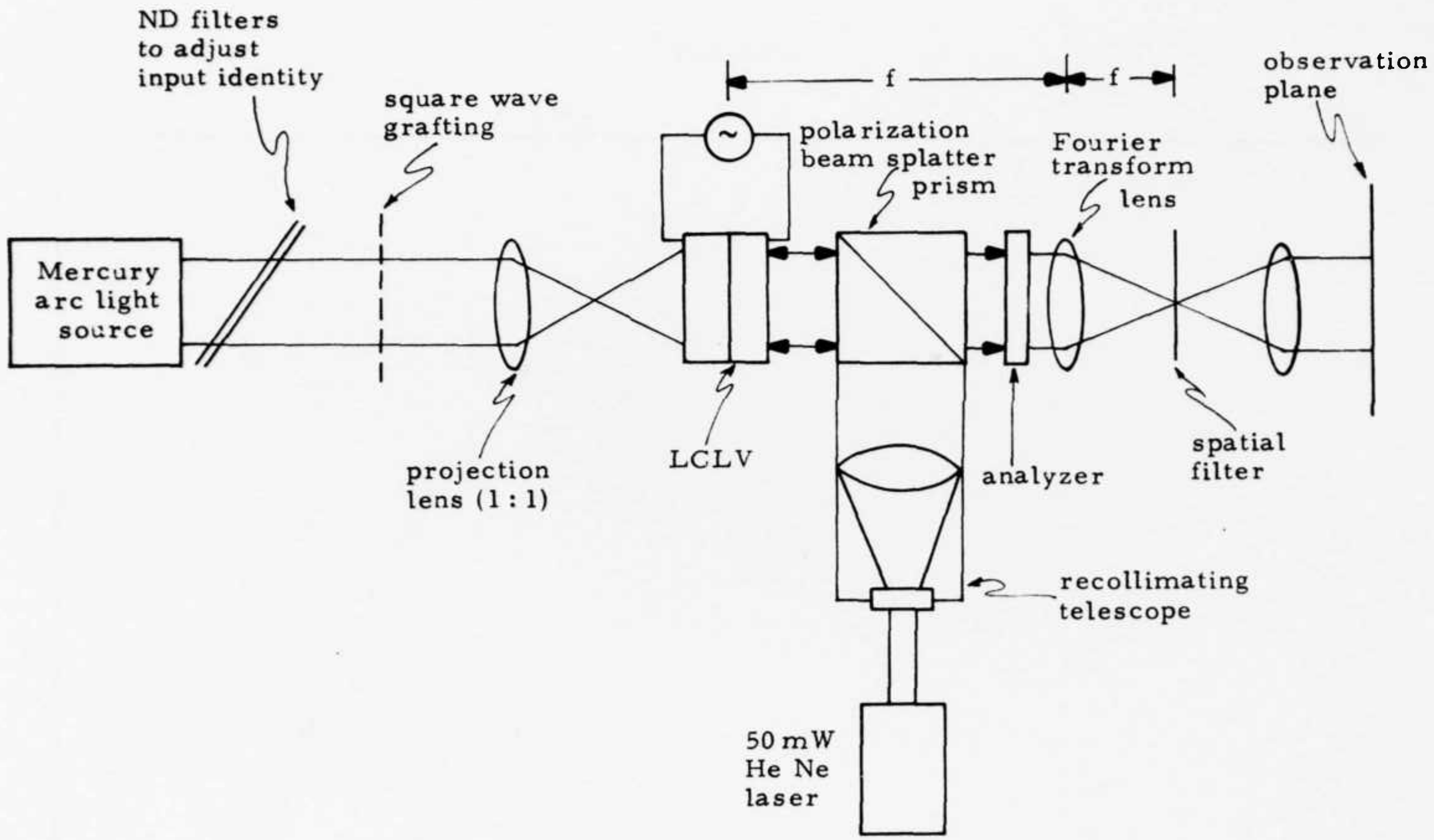


Figure 1. Experimental Setup

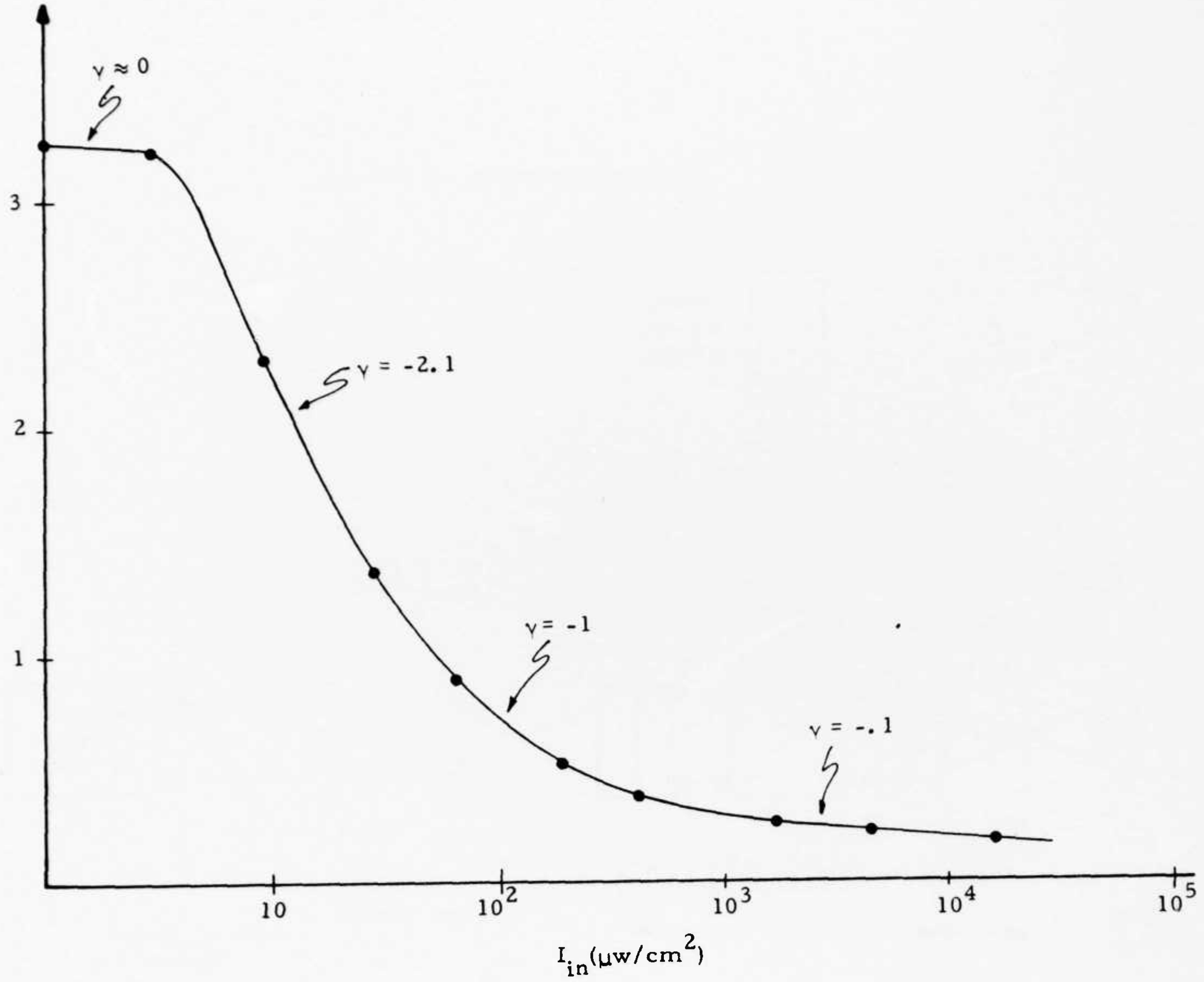


Figure 2. Densitometry Curve V = 4.0 volts rms at 2 KHz.

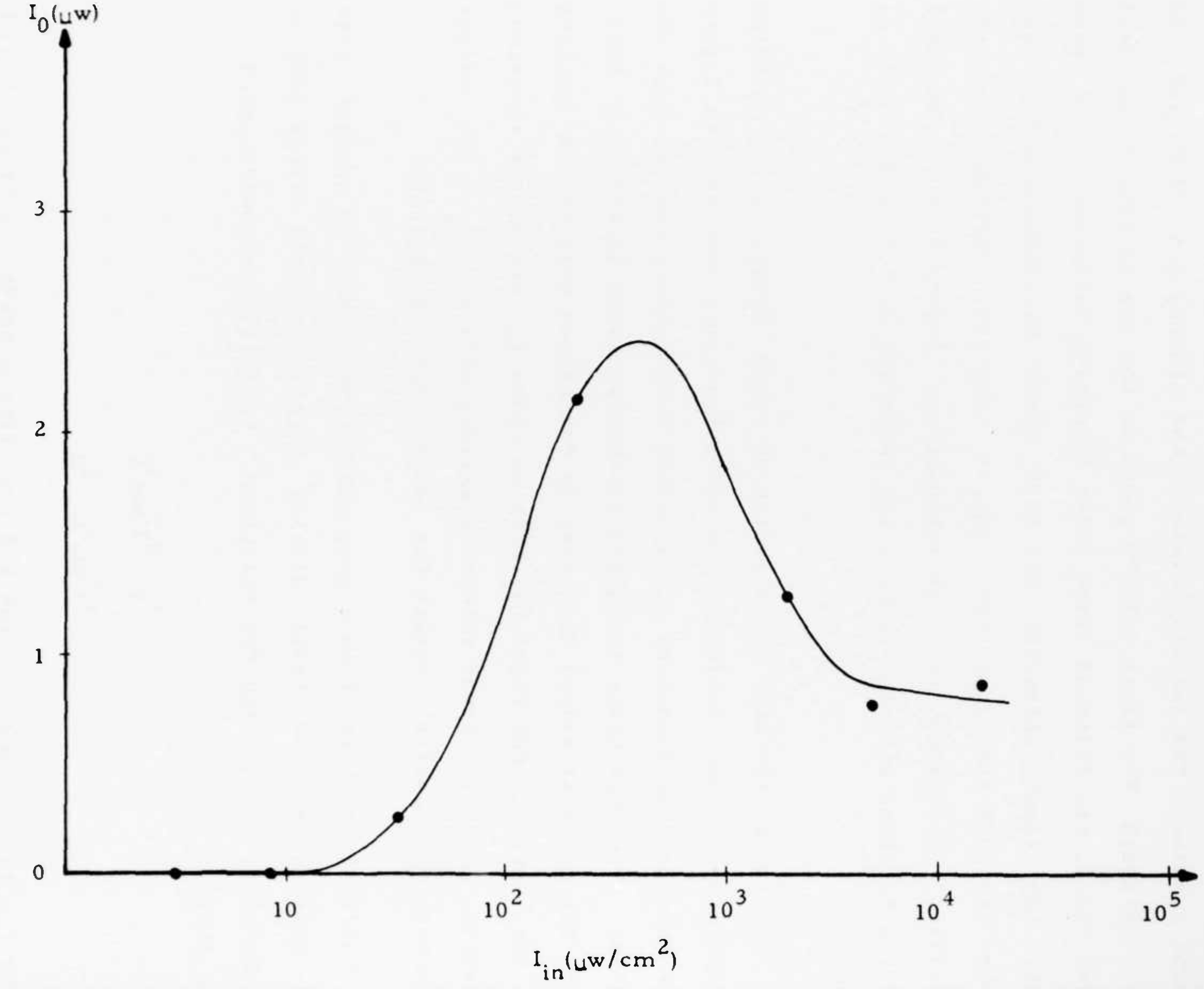


Figure 3 - Level Slicing I/O Characteristic for V = 4.0 volts rms at 2KHz.

experiment. The input transparency is sampled by a 12 cycles/mm Ronchi grating. The density levels of the grating are .21 ND and .68 ND. Although the light valve's gamma is too low to achieve an ideal level slice, the filtered first order intensity increases to a peak where the input intensity at high gamma has maximum modulation. Beyond the peak the first order output intensity decreases as the modulation decreases at high intensities. Figure 4 shows the level slicing function with the light valve operating at 5.8 volts RMS at 5 KHz.

We studied the LCLV characteristics that impact level slicing performance. The results can be understood with reference to figure 5, which shows an idealized input/output relationship for the LCLV ODP system. The LCLV has a threshold incoherent input intensity I below which the coherent output intensity is determined only by the quality of the optics. For input intensities above I_T, the output increases linearly. When the input reaches a maximum intensity I_n, the output saturates and remains constant for larger input intensities.

When the LCIV input is a gray scale transparency in contact with the grating of the level slicing experiments, the output will be modulated by the grating for tranparency transmission coefficients T_T such that

$$I_T < T_T T_{max} I_L$$

and

$$T_T T_{min} I_L < I_M$$

where T_{max} and T_{min} are the high and low transmission coefficients of

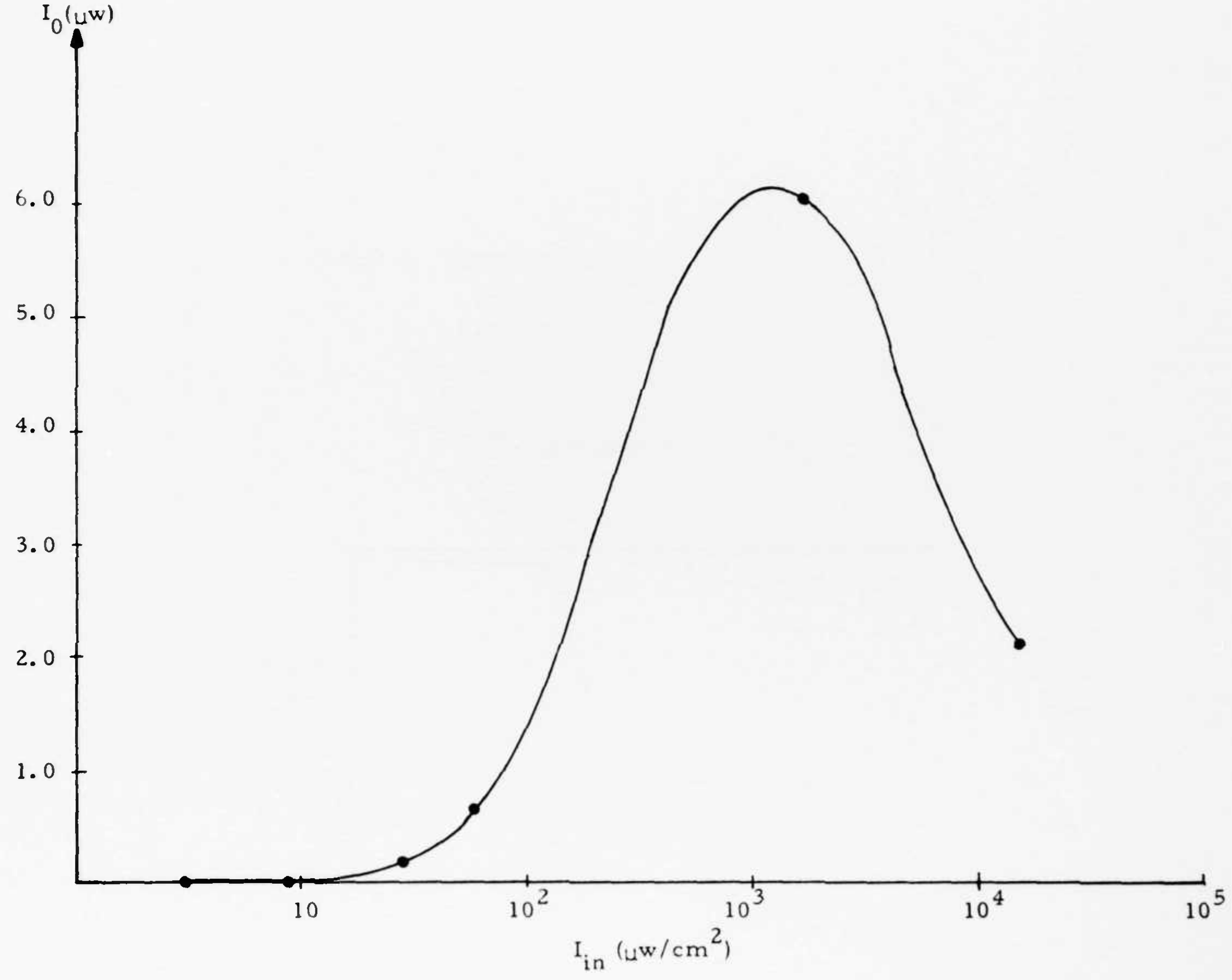


Figure 4. Level Slicing I/O Characteristic for V = 5.8 volst rms at SKHz

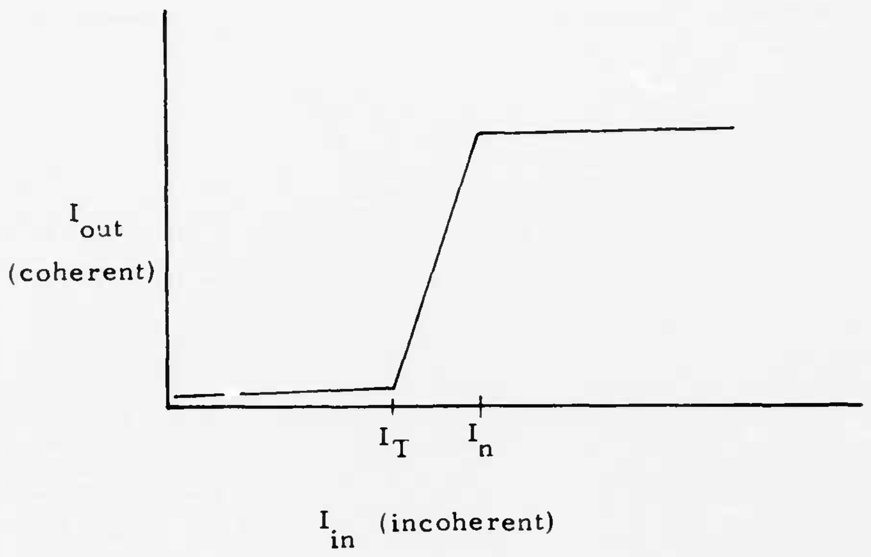


Figure 5. Idealized LCLV I/O Characteristic

the grating and I_L is the incident lamp intensity. Replacing the inequalities by equalities, the gray scale (density) range that appears as a single "sliced" level is

$$\Delta D_S = \log \frac{I_M}{I_T} \frac{T_{max}}{T_{min}} .$$

The total number of levels that can be independently sliced out depends on the maximum lamp intensity, I_{Lmax}. The transparency density range covered is

$$D_{total} = \log \frac{T_{min} I_{L\,max}}{I_M} ,$$

and the number of levels, n, is

$$n = \frac{D_{total}}{\Delta D_S} .$$

These equations can be expressed in terms of quantities that are characteristic of the system components as follows. The differential density of the grating is

$$\Delta D_G = \log \frac{T_{max}}{T_{min}} .$$

The device "density" range is

$$\Delta D_D = \log \frac{I_M}{I_T} ,$$

which is, in a photographic analogy, the number of "log exposure" units required to go from threshold to saturation. The lamp "density" is

$$D_L = \log \frac{I_{L\,max}}{I_M} ,$$

i.e., the number of saturation "log exposure" units available from the lamp. Finally, the basic grating density is

$$D_G = - \log T_{min} .$$

With these definitions, we have

$$\Delta D_S = \Delta D_D + \Delta D_G$$

for the width of the slice, and

$$D_{total} = D_L - D_G$$

for the total density range that can be sliced. The number of level slices in this range is

$$n = \frac{D_L - D_G}{\Delta D_D + \Delta D_G} .$$

Therefore, in this simple picture, the number of level slices can be maximized by using a low density grating with a small density step, a bright lamp and a LCLV with a small ΔD_D, i.e., a small I_M/I_T. To some extent, the system can be improved by increasing the device threshold, I_T, if the intensity range $I_M - I_T$ is constant. However, this also increases I_M and therefore decreases the available density range D_{total}.

The threshold intensity, I_T, is determined primarily by the "switching ratio" of the LCLV photoconductor (see earlier LCLV reference). For linear operation, a low threshold is desired, and consequently a low switching ratio is used. This is the type of device used in our preliminary experiments.

A higher switching ratio is obtained with a thicker CdS photoconductor layer in the LCLV. Such devices are now being fabricated and an improvement from ~ 1 level to 3-5 levels is indicated. Thus far we have not been able to obtain a thick film device for the nonlinear ODP experiments, though attempts are continuing.

Work is also underway (on another program) to develop silicon photoconductor LCLVs. The silicon characteristics indicate that up to 10 levels could be isolated with the same equipment and a silicon LCLV. It is unlikely, however, that a silicon device will be available for tests.

The most lucrative device modification is to find liquid crystal materials that give a smaller $I_M - I_T$. This would reduce ΔD_D and increase D_L, both of which would increase the number of levels n. There is a good probability that addition of some cholesteric liquid crystal to the device mixture would provide some improvement in I_M. However, the required liquid crystal chemistry is non-trivial and experiments of this kind are beyond the scope of the current program.

We are in the process of making some noise measurements on the LCLV, but the results are preliminary and will not be reported here.

5. RECENT PH.D DISSERTATIONS

This section includes the abstracts of recent Ph.D. lissertations completed over the past six-month research period. Due to the length of their reports, it was felt that abstracts presented in this semi-annual report was appropriate. For those individuals interested in further details, these dissertaticns appear in their entirety in the form of USC-Image Processing Institute reports. They are available upon request.

5.1 NONLINEAR OPTICAL IMAGE PROCESSING WITH HALFTONE SCREENS

Stephen R. Dashiell

Coherent optical systems are of interest in image or data processing because of their ability to rapidly handle large bandwidth data in parallel. They have been restricted to performing linear operations such as Fourier transformation and convolution, due to the inherent linear nature of an optical system at low power levels. In this dissertation, the combination of a nonlinear halftoning step followed by band-pass spatial filtering to yield a specific nonlinear intensity transfer function is explored.

A general analysis of the problem assuming infinite copy film gamma and saturation density is made. A constructive algorithm for designing a halftone cell shape and selecting the diffraction order to yield very general types of nonmonotonic nonlinearities is presented. Numerous examples of the synthesis procedure are given.

The design of nonmonotonic halftone cells which allow a nonmonotonic nonlinearity with an arbitrary number of changes of sign in slope to be obtained in the first diffraction order is considered. An iterative algorithm suitable for computer implementation, and numerous examples of halftone cells designed with this algorithm are given.

The effects of allowing the film gamma and saturation density to become finite are analyzed, and a technique for compensating a priori for some of the resulting degradations is given.

Experimental results with general halftone screens made on a plotting flatbed microdensitometer are presented. Logarithmic, exponential, and level slice characteristics have been achieved with monotonic cells. Intensity notch filter and quantization characteristics have been achieved with nonmonotonic cells. Other generalizations of the technique are discussed, including the possibility of real-time nonlinear processing with optical input transducers.

5.2 ESTIMATION AND DETECTION OF IMAGES DEGRADED BY FILM-GRAIN NOISE

Firouz Naderi

Film-grain noise is a term describing the intrinsic noise produced by a photographic emulsion during the process of recording an image on film. Although film-grain noise , s recently been considered within the field of image processing, the nature of the noise is still somewhat misunderstood.

One goal of this study has been to use the theoretical and experimental results on film characteristics obtained by photographic scientists in order to define film-grain noise within the context of estimation theory. A detailed model for the overall photographic imaging system is presented. There are linear blurring effects at the initial and the final segments of this model to account for various optical and chemical degradations. The middle segment of the model represents signal dependence effects of film-grain noise and includes a nonlinear noise term. The accuracy of this model is tested by simulating images according to it and comparing the results to images of similar targets that were actually recorded on film. These comparisons point out that the model is a gcod representation of the photographic imaging system.

The restoration of images degraded by film-grain noise is considered in two different contexts - estimation theory and detection theory. Under the topic of estimation, a discrete wiener filter is developed which explicitly allows for the signal-dependence of the noise. The filter adaptively alters its characteristics based on non-stationary first order statistics of an image. This filter is shown to have an advantage over the conventional wiener filter.

In the case of extremely low contrast images digitized by a very small aperture, film-grain noise is so severe that conventional statistical restoration techniques have little effect. For use in this situation, a heuristic algorithm is developed which incorporates some of the vision properties of the human observer. Bayesian detection theory is used to justify the procedure and to provide some insight into its use. This algorithm also explicitly includes the signal dependence on the noise and has the capability of greatly outperforming the human observer in locating objects corrupted by very severe noise.

Experimental results for modeling, the adaptive estimation filter and the Bayesian detection algorithm are presented.

5.3 IMAGE RESTORATION BY SPLINE FUNCTIONS

Mohammad J. Peyrovian

Spline functions, because of their highly desirable interpolating and approximating characteristics, are used as a potential alternative to the conventional pulse approximation method in digital image processing. For uniformly spaced knots, a class of spline functions called B-splines has the useful properties of shift invariance, positiveness, convolutional property and local basis property. These properties are exploited in image processing for linear incoherent imaging systems.

The problem of image degradation in a linear imaging system is described by a superposition integral. For simulation of degradation and restoration by means of a digital computer, the continuous imaging model must be discretized. Thus, a theoretical and experimental study of quadrature formulae, particularly monospline and best quadrature formulae in the sense of Sard, is presented. It is shown that a good choice of degree for a monospline highly depends on the frequency content of the integrand, and in most cases, a cubic monospline generates less error than the pulse approximation method and Newton-Cotes quadrature formulae.

In space-invariant imaging systems, the object and point-spread function are represented by B-splines of degrees m and n. Exploiting the convolutional property, the deterministic part of the blurred image is a spline function of degree m+n+1. A minimum norm principle leading to pseudo-inversion is used for the restoration of space variant degradations and underdetermined and overdetermined models. Space-variant point-spread functions that describe astigmatism and curvature of field are derived and coordinate transformations are applied to reduce the dimensionality. The singular-value-decomposition technique is used for solution of the simplified equations.

For noisy blurred images, a controllable smoothing criteria based on the locally variable statistics and minimization of the second derivative is defined, and the corresponding filter, applicable to both space-variant and space-invariant degradations, is obtained. The parameters of the filter determine the local smoothing window and overall extent of smoothing, and thus the trade-off between resolution and smoothing is controllable in a spatially nonstationary manner. Since the matrices of this filter are banded circulant or Toeplitz, efficient algorithms are used for matrix manipulations.

5.4 INTERFRAME CODING OF DIGITAL IMAGES USING TRANSFORM AND HYBRID TRANSFORM/PREDICTIVE TECHNIQUES

John A. Roese

In the design of digital image coding systems, the principal objective is to achieve high quality receiver image reconstructions with a minimum number of transmitted code bits. Bit rate reductions

are achieved by exploiting statistical redundancies within an image. This is combined by transmission of only those portions of the mathematical image representation which the human observer is most sensitive to. This dissertation describes research intended to extend current image coding techniques to the coding of sequences of digital images transmitted over a digital communications channel. The emphasis is directed towards definition of an image coding system that exploits temporal as well as spatial image redundancies.

A primary objective of this investigation is to develop a class of interframe hybrid transform/predictive coders having near optimum levels of performance. The interframe hybrid coder implementations considered employ two-dimensional unitary transforms in the spatial domain coupled with first-order DPCM predictive coding in the temporal domain. Based on a statistical image representation, a model is developed for the hybrid coder transform coefficient temporal difference variance matrix. With this model, theoretical MSE performance levels for the hybrid coder with zonal coding are determined as a function of spatial subblock size.

Implementations of the interframe hybrid coder using discrete cosine and Fourier transforms are experimentally evaluated. High quality image reconstruction are demonstrated for reductions of 32:1 in average pixel bit rate. Operational considerations investigated for the hybrid interframe coder include initial conditions, spatial and temporal adaptation, reinitialization, and total transmission bit rate. Also, sensitivity of the interframe hybrid coder to channel error is studied.

Comparisons are drawn between hybrid transform/predictive and three-dimensional transform interframe coders. Theoretical zonal sampling and zonal coding MSE performance for three-dimensional cosine transform coders are evaluated for different frame storage requirements and spatial subblock sizes.

A tabular summary of experimental performance and system design parameters for the main classes of interframe coders is presented.

6. PUBLICATIONS

1. H.C. Andrews, "A Potpourri of Mathematical Techniques Applied to Image Science," Proceedings of the First Annual Symposium on Mathematical Techniques in Image Science, USNPGS, Monterey, California, November 1976.

2. H.C. Andrews, "Image Processing at the University of Southern California," CalTech/JPL Conference on Image Processing, November 1976.

3. H.C. Andrews and B.R. Hunt, Digital Image Restoration, Prentice Hall, New Jersey, 1977.

4. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Analysis and Synthesis," submitted to Applied Optics.

5. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effects of Input Medium and Precompensation," submitted to Applied Optics.

6. S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Non-monotonic Halftone Cells and Phase Halftones," submitted to Applied Optics.

7. W. Frei, "Image Enhancement by Histogram Hyperbolization," accepted for publication in Computer Graphics and Image Processing.

8. W. Frei, "Rate-Distortion Coding Simulation for Color Images," Proceedings of the SPIE, San Diego, California, August 1976.

9. H.S. Hou and H.C. Andrews, "Least-Squares Image Restoration Using Spline Basis Functions," accepted for publication in IEEE Transactions on Computers.

10. D.G. McCaughey and H.C. Andrews, "Degrees of Freedom for Projection Imaging," accepted for publication in IEEE Transactions on Acoustics, Speech, and Signal Processing.

11. N.E. Nahi and M.H. Jahanshahi, "Image Boundary Estimation," accepted for publication in IEEE Transactions on Computers.

12. N.E. Nahi, "Estimation of Object Boundaries in Noisy Images," Proceedings of 1976 IEEE Decision and Control Conference, Clear Water, Florida, December 1976, (invited paper).

13. N.E. Nahi, "Image Modelling by Replacement Processes," Symposium on Current Mathematical Problems in Image Science, November 1976, (invited paper).

14. R. Nevatia, "Depth Measurement by Motion Stereo," Computer Graphics and Image Processing, June 1976, pp. 203-214.

15. R. Nevatia, "A Color Edge Detector," to appear in the Proceedings of the Third International Joint Conference on Pattern Recognition.

16. R. Nevatia, "Characteristics and Requirements of Computer Vision Systems," to be presented at the Workshop on Computer Vision Systems, University of Massachusetts, June 1977.

17. R. Nevatia and T.O. Binford, "Description and Recognition of Curved Objects," accepted for publication in the Artificial Intelligence Journal.

18. W.K. Pratt, "Digital Image Enhancement and Restoration Display Techniques," Society for Information Display, 1976 SID International Symposium, Los Angeles, California, May 1976.

19. W.K. Pratt, "Pseudoinverse Image Restoration Computational Algorithms," presented at 1976 Joint Soviet-USA Optical Information Processing Symposium, Novasibirsk Siberia: to be published in Optical Information Processing, Vol. II, Plenum Press, 1977.

20. W.K. Pratt, "Two-Dimensional Unitary Transforms," presented at NATO ASI Symposium on Digital Image Processing: to be published in Proceedings NATO ASI Symposium on Digital Image Processing, Plenum Press, 1977.

21. J.A. Roese and W.K. Pratt, "Theoretical Performance Models for Interframe Transform and Hybrid Transform/DPCM Coders," Proceedings SPIE Symposium on Advances in Image Transmission Techniques, San Diego, California, August 1976.

22. A.A. Sawchuk, "Signal-Dependent Noise Models and Nonlinear Nonstationary Image Restoration," NATO Advanced Study Institute on Digital Image Processing and Analysis, Bonas, France, June 1976.

23. A.A. Sawchuk, "Nonlinear Noumonotonic Optical Processing," Gordon Research Conference on Coherent Optics and Holography, Santa

Barbara, June 1976, (invited paper).

24. A.A. Sawchuk, "Image Restoration and Enhancement with Signal-Dependent Noise Models," Engineering Foundation Conference - Algorithms for Image Processing, Rindge, New Hampshire, August 1976, (invited paper).

25. A.A. Sawchuk, "Artificial Stereo," Engineering Foundation Conference - Algorithms for Image Processing, Rindge, New Hampshire, August 1976, (invited paper).

26. A.A. Sawchuk and S.R. Dashiell, "Nonlinear Coherent Optical Processing - Synthesis Algorithms and Real-Time Systems," Proceedings Society of Photo-Optical Instrumentation Engineers Technical Symposium - Optical Information Processing, Vol. 83, San Diego, August 1976.

27. A.A. Sawchuk, Session Organizer and Co-chairman, Optical Information Processing: Real-Time Devices and Novel Techniques Seminar, Society of Photo-Optical Instrumentation Engineers Technical Symposium, San Diego, August 1976; Editor, Vol. 83, SPIE Proceedings.